



(19)

Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 1 130 849 A2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
05.09.2001 Bulletin 2001/36

(51) Int Cl.⁷: H04L 12/26, H04L 12/56

(21) Application number: 01104498.9

(22) Date of filing: 01.03.2001

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 01.03.2000 JP 2000056254
11.12.2000 JP 2000376615

(71) Applicant: FUJITSU LIMITED
Kawasaki-shi, Kanagawa 211-8588 (JP)

(72) Inventors:
• Soumiya, Toshio
Nakahara-ku, Kawasaki-shi, Kanagawa 211 (JP)
• Nakamichi, Koji
Nakahara-ku, Kawasaki-shi, Kanagawa 211 (JP)
• Takashima, Kenya
Nakahara-ku, Kawasaki-shi, Kanagawa 211 (JP)

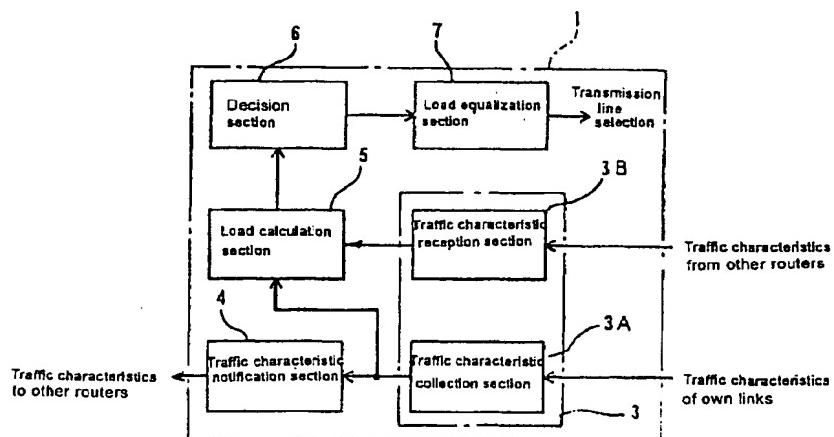
(74) Representative: HOFFMANN - EITLE
Patent- und Rechtsanwälte
Arabellastrasse 4
81925 München (DE)

(54) Transmission path controlling apparatus, method and program

(57) The present invention relates to the distribution of load in a plurality of routes from a communication apparatus which serves as a start point to another communication apparatus which serves as an end point from among communication apparatus in a network so that load distribution of load can be performed between the set routes. An apparatus provided in a router in a network includes a traffic characteristic collection section for collecting traffic characteristics of transmission paths

connected to the router, a traffic characteristic notification section notifies other routers of the collected traffic characteristics, a load calculation section calculates load based on the collected traffic characteristics, a decision section decides based on the determined load information whether or not a transmission path should be added or deleted, and a load equalization section equalizes the determined load among the plurality of transmission paths.

FIG.1(b)



BEST AVAILABLE COPY

EP 1 130 849 A2

Description

BACKGROUND OF THE INVENTION

5 **Field of the Invention**

[0001] This invention relates to a transmission path controlling apparatus provided in a communication apparatus (for example, a router or load equalization apparatus) which forms a network, and further to a transmission path controlling method executed by the apparatus as well as a medium on which a transmission path controlling program used in the apparatus is recorded. The method and apparatus may be provided within, for example, an Internet Protocol network (in the following description an Internet Protocol is sometimes referred to simply as IP; accordingly, an Internet Protocol network is referred to simply as IP network).

10 **Description of the Related Art**

[0002] The Internet is intended to handle all communication applications on the internet protocol and is being popularized rapidly at present. Further, the Internet originally has a connectionless network architecture which does not involve establishment of a connection, and is routed to an outgoing path based on a destination address described in an IP packet. In short, in a communication apparatus having the function described, at a point of time when an IP packet arrives at the apparatus, the IP packet is merely transferred to an outgoing path which corresponds to the IP address.

[0003] In the Internet having such a network architecture as described above, in order to determine to which outgoing path each IP address should be transferred, a protocol for path determination called routing protocol is handled between communication apparatus.

[0004] At present, as an algorithm for the routing protocol, a system which employs the algorithm of Dijkstra is commonly used. However, the routing protocol mentioned has a problem in that a plurality of routes to a destination cannot be selected or the routing protocol is executed only when the network topology varies.

[0005] In short, with such a technique as described above, once a route is established, an IP packet is transferred only along the route, and as a result, there is the possibility that chronic congestion occurs.

[0006] Further, such a technique as described above has a problem also in that, since a plurality of routes cannot be set, when a certain route suffers from congestion, even if some other route is unoccupied, this unoccupied route cannot be used.

[0007] Thus, the following techniques have been proposed.

35 (1-1) Equal cost multipath technique in the OSPF (Open Shortest Path First)

(1-2) Multipath technique in the IBGP (Internal Border Gateway Protocol)

[0008] In particular, the method according to the technique of (1-1) is a method which uses a function of the OSPF, and uses a function capable of laying a plurality of paths having an equal cost from a communication apparatus which serves as a start point to another communication apparatus which serves as an end point to distribute load among the plurality of paths (multipaths).

[0009] Meanwhile, the method according to the technique of (1-2) uses a function of the IBGP to lay a plurality of paths and distribute load among the plurality of paths (multipaths).

[0010] However, the method which uses a function of the OSPF according to (1-1) above has the following problems.

[0011] In particular, there is a restriction in that the multipaths must be equal in cost, and the degree of freedom in path selection is low. Further, since the OSPF is used, the method cannot cope with a variation of the network load on a real time basis.

[0012] Also the method which uses a function of the IBGP according to (1-2) above has a restriction in that the multipaths must be equal in cost and has a problem in that the degree of freedom in path selection is low, similarly to the problems of the method which uses a function of the OSPF.

[0013] It is to be noted that also a technique has been proposed wherein a path is set in the layer 3 between data processing apparatus which connect networks to each other and load is distributed in a unit of a route in the layer 3 to dynamically provide a traffic balance to improve the network performance (refer to the official gazette of Japanese Patent Laid-Open No. HEI 10-224400). With the technique, however, even if a traffic balance can be provided between networks, load distribution in each network cannot be performed. Further, only load distribution in the layer 3 and in a unit of a route is taken into consideration, and there is the possibility that load distribution control may be complicated.

SUMMARY OF THE INVENTION

[0014] The present invention has been made in view of such problems as described above, and it is an object of the present invention to provide a transmission path controlling apparatus and a transmission path controlling method as well as a medium on which a transmission path controlling program is recorded, wherein a plurality of routes (transmission paths) from a communication apparatus (source communication apparatus) which serves as a start point to another communication apparatus (destination communication apparatus) which serves as an end point from among communication apparatus in a network such that load can be distributed among the set routes so as to allow traffic engineering in a network such as the Internet irrespective of the network topology and without depending upon the type of the transmission paths.

[0015] FIGS. 1(a), 1(b) and 1(c) are explanatory views of a principle of the present invention. First, a transmission path controlling apparatus according to the present invention is used with such a network as shown in FIG. 1(a), the network includes a source communication apparatus (an origination communication apparatus) 1S, a destination communication apparatus 1D, a plurality of transmission paths 2-i (i = 1 to N: N is a natural number) capable of being set between the source communication apparatus 1S and the destination communication apparatus 1D, and an intermediate communication apparatus 1R interposed in any one of the transmission paths 2-i, and provided in one of the communication apparatus 1S, 1D and 1R which form the network (in the following description, where the source communication apparatus 1S, destination communication apparatus 1D and intermediate communication apparatus 1R are represented without specific distinction among them, they are referred to merely as communication apparatus 1, and the network is an IP network for example).

[0016] The transmission path controlling apparatus (the source communication apparatus 1S and the destination communication apparatus 1D) according to the present invention includes, as shown in FIG. 1(b), a traffic characteristic collection section 3, a traffic characteristic notification section 4, a load calculation section 5, a decision section 6 and a load equalization section 7.

[0017] Here, the traffic characteristic collection section 3 collects traffic characteristics of those of the transmission paths 2-i connected to the communication apparatus 1 or the other communication apparatus. The traffic characteristic collection section 3 includes a traffic characteristic collection section 3A for collecting traffic characteristics at the communication apparatus 1 itself, and a traffic characteristic reception section 3B for receiving traffic characteristics notified from the other communication apparatus.

[0018] The traffic characteristic notification section 4 notifies the other communication apparatus of the traffic characteristics collected by the traffic characteristic collection section 3 (particularly 3A).

[0019] The load calculation section 5 calculates load based on the traffic characteristics collected by the traffic characteristic collection section 3.

[0020] The decision section 6 discriminates whether or not a transmission path should be added or deleted based on the load information determined by the load calculation section 5, and the load equalization section 7 equalizes the load determined by the load calculation section 5 among the plurality of transmission paths.

[0021] The transmission path controlling apparatus (the intermediate communication apparatus 1R) according to the present invention includes, as shown in FIG. 1(c), a traffic characteristic collection section 3 and a traffic characteristic notification section 4.

[0022] Here, the traffic characteristic collection section 3 collects traffic characteristics of those of the transmission paths 2-i connected to the communication apparatus 1 or the other communication apparatus. The traffic characteristic collection section 3 includes a traffic characteristic collection section 3A for collecting traffic characteristics at the communication apparatus 1 itself, and a traffic characteristic reception section 3B for receiving traffic characteristics notified from the other communication apparatus.

[0023] The traffic characteristic notification section 4 notifies the other communication apparatus of the traffic characteristics collected by the traffic characteristic collection section 3 (3A and 3B).

[0024] It is to be noted that, while it is possible to provide the transmission path controlling apparatus, which is provided in each of the source communication apparatus 1S, destination communication apparatus 1D and intermediate communication apparatus 1R, with such a traffic characteristic collection section 3, traffic characteristic notification section 4, load calculation section 5, decision section 6 and load equalization section 7 as described above, it is possible to provide the transmission path controlling apparatus, which is provided in the source communication apparatus 1S, with the traffic characteristic collection section 3, traffic characteristic notification section 4, load calculation section 5, decision section 6 and load equalization section 7 and provide, as shown in figure 1(c) the transmission path controlling apparatus, which is provided in each of the destination communication apparatus 1D and the intermediate communication apparatus 1R other than the source communication apparatus 1S, only with the traffic characteristic collection section 3 and traffic characteristic notification section 4.

[0025] It is to be noted that the traffic characteristic collection section 3 may have means for smoothing the collected traffic characteristic values, or may have means for determining a maximum value of the collected traffic characteristic

values as a representative value of the traffic characteristic values.

[0026] Further, the traffic characteristic notification section 4 may have means for notifying (providing) the information regarding the traffic characteristics making use of a packet provided by a routing protocol, or may have means for notifying the information regarding the traffic characteristics after each of a fixed period, or else may have means for notifying the information regarding the traffic characteristic with a message of the RSVP expanded.

[0027] Further, the traffic characteristic collection section 3 may have means for collecting a unit of a flux of flows for collecting traffic characteristics in a unit of a label which is used by a communication apparatus of the label switch type such as MPLS (Multi-Protocol Label Switching).

[0028] The traffic characteristic collection section 3 provided in the source communication apparatus 1S may have means for discriminating a utilization (a utilization rate) of each of the transmission paths 2-i connected to the source communication apparatus 1S based on an average utilization (an average utilization rate) of the transmission paths 2-i connected to the intermediate communication apparatus 1R collected from the intermediate communication apparatus 1R.

[0029] It is to be noted that the traffic characteristic collection section 3 provided in the source communication apparatus 1S may have means for discriminating a utilization of each of the transmission paths 2-i connected to the source communication apparatus 1S based on a maximum utilization of the transmission paths 2-i connected to the intermediate communication apparatus 1R collected from the intermediate communication apparatus 1R.

[0030] Further, the traffic characteristic collection section 3 provided in the source communication apparatus 1S may have means for discriminating after each fixed period whether or not each of the transmission paths 2-i is in congestion.

[0031] The load calculation section 5 provided in the source communication apparatus 1S may have means for calculating effective load taking a number of abandoned packets generated in the intermediate communication apparatus 1R into consideration. In this instance, an effective load calculation unit may be constructed so as to set an upper limit to the effective load when the effective load is to be calculated.

[0032] Furthermore, the transmission path controlling apparatus may be constructed such that the load calculation section 5 provided in the source communication apparatus 1S has means for calculating effective load taking a number of abandoned packets generated in the intermediate communication apparatus 1R into consideration, and the decision section 6 provided in the source communication apparatus 1S has means for calculating an effective bandwidth from the effective load obtained by the effective load calculation unit and determining, regarding all of the transmission paths as one virtual pipe, a utilization of the virtual pipe.

[0033] Further, the transmission path controlling apparatus may be constructed such that the load calculation section 5 provided in the source communication apparatus 1S has means for calculating effective load taking a number of abandoned packets generated in the intermediate communication apparatus 1R into consideration, and the decision section 5 provided in the source communication apparatus 1S has a pipe utilization calculation unit for calculating an effective bandwidth from the effective load calculated by the effective load calculation unit and determining, regarding all of the transmission paths 2-i as one virtual pipe, a utilization of the virtual pipe, and comparing the pipe utilization obtained by the pipe utilization calculation unit with a set reference value and determining whether a transmission path is to be added or deleted in accordance with a result of the comparison.

[0034] The decision section 6 provided in the source communication apparatus 1S may be constructed so as to add a transmission path when the pipe utilization obtained by the pipe utilization calculation unit is higher than a reference value for addition of a path, or may be constructed so as to delete a transmission path having a low load when the pipe utilization obtained by the pipe utilization calculation unit is lower than a reference value for deletion of a path.

[0035] The load equalization section 7 provided in the source communication apparatus 1S may have a move effective bandwidth calculation unit for determining an effective bandwidth to be moved with each of the transmission paths 2-i from an average effective bandwidth calculated for all of the transmission paths 2-i based on an effective bandwidth obtained from effective load obtained taking a number of abandoned packets generated in the intermediate communication apparatus 1R into consideration.

[0036] Further, in this instance, the load equalization section 7 provided in the source communication apparatus 1S may distribute packets flowing into the source communication apparatus 1S to the plurality of transmission paths in proportion to the effective bandwidths to be moved calculated by the move effective band calculation unit.

[0037] It is to be noted that the load equalization section 7 provided in the source communication apparatus 1S may be constructed so as to perform a calculation using a Hash function based on an address and distribute traffic flows at random based on a result of the calculation.

[0038] Then, in the present invention, in a network which includes a source communication apparatus 1S, a destination communication apparatus 1D, a plurality of transmission paths 2-i capable of being set between the source communication apparatus 1S and the destination communication apparatus 1D, and an intermediate communication apparatus 1R interposed in any one of the transmission paths 2-i,

(A1) a traffic characteristic collection step of collecting traffic characteristics of those of the transmission paths 2-i

connected to the other communication apparatus, and
(A2) a traffic characteristic notification step of notifying the source communication apparatus 1S of the traffic characteristics collected by the traffic characteristic collection step are performed by communication apparatus other than the source communication apparatus of the network, and
5 (B1) a traffic characteristic collection step of collecting traffic characteristics of those of the transmission paths connected to the source communication apparatus 1S,
(B2) a load calculation step of calculating load based on one or both of the traffic characteristics collected by the traffic characteristic collection step and the traffic characteristics obtained from the other communication apparatus,
10 (B3) a determining step of determining whether or not a transmission path should be added or deleted based on the load information determined by the load calculation step, and
(B4) a load equalization step of equalizing the load determined by the load calculation step among the plurality of transmission paths may be performed by the source communication apparatus 1S.

15 [0039] Also a medium is disclosed on which a transmission path controlling program for being used by one of the communication apparatus 1 which forms the network shown in FIG. 1(a) is recorded, the transmission path controlling program causing a computer to function as a traffic characteristic collection unit for collecting traffic characteristics of those of the transmission paths connected to the communication apparatus or the other communication apparatus, traffic characteristic notification unit for notifying the other communication apparatus of the traffic characteristics collected by the traffic characteristic collection unit, load calculation unit for calculating load based on the traffic characteristics collected by the traffic characteristic collection unit, decision section for deciding whether or not a transmission path should be added or deleted based on the load information determined by the load calculation unit, and a load equalization unit for equalizing the load determined by the load calculation unit among the plurality of transmission paths.
20

25 BRIEF DESCRIPTION OF THE DRAWINGS

[0040]

30 FIG. 1(a) is an example of a network of the present invention;
FIG. 1(b) is an example of an embodiment of the present invention;
FIG. 1(c) is an example of another embodiment of the present invention.
FIGS. 2(a) and 2(b) are views for explaining an MPLS network.
FIG. 3 is a view for explaining a label used in MPLS.
FIG. 4 is a view for explaining the label used in the MPLS.
35 FIG. 5 is a view for explaining an LSP.
FIG. 6 is a view for explaining a flow of a message of the RSVP-LSP-Tunnel.
FIG. 7 is a view for explaining an MPLS network applied to an embodiment of the present invention.
FIG. 8 is a functional block diagram of an MPLS router as an embodiment of the present invention.
FIG. 9 is a functional block diagram showing details of a traffic engineering section of the MPLS router as the 40 embodiment of the present invention.
FIG. 10 is a view for explaining an effective load.
FIG. 11 is a view for explaining an effective bandwidth to be moved.
FIG. 12 is a view for explaining an example of a manner of load adjustment.
FIG. 13 is a view for explaining an example of a Hash function.
45 FIG. 14 is a view for explaining a view of a network for experiment of load adjustment.
FIG. 15 is a view for explaining an example of a result of evaluation of a load adjustment performance.
FIG. 16 is a view for explaining another example of a result of evaluation of a load adjustment performance.

50 DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0041] In the following, an embodiment of the present invention is described with reference to the drawings.
[0042] As an embodiment of the present invention, a case wherein an MPLS router as a router (an example of a communication apparatus) of the label switch type is equipped with the present transmission path controlling apparatus and a route different from a routing protocol is set using the MPLS (Multi-Protocol Label Switching) and using a label 55 function of the MPLS to effect load distribution is described. Here, the MPLS signifies a system which can simply and effectively perform a service which is difficult to realize with a router at present by using a label of the layer 2 in a specialized manner for an IP.
[0043] In the following, an outline of the MPLS is described first, and then a load distribution method and a load

distribution algorithm (including description of operation) of the MPLS are described.

(2-1) Description of an Outline of the MPLS

- 5 [0044] First, an outline of the MPLS is described. FIGS. 2(a) and 2(b) show a network architecture and a manner of basic packet transfer of the MPLS. An MPLS network is composed of an edge label switching router (Edge Label Switching Router, hereinafter referred to as edge LSR) 11E and a core label switching router (Core Label Switching Router, hereinafter referred to as core LSR) 11C. This is called MPLS domain.
10 [0045] Here, the edge LSR 11E is a router which is positioned on the boundary with an existing internet network and adds a label to a packet to realize a network layer service of a high performance and a high added-value. This edge LSR terminates a LAN interface of an unlabeled frame base (Frame Base).
[0046] Meanwhile, the core LSR 11C is a router which exchanges a label of a labeled packet or cell. This core LSR 11C may be constructed so as to support not only label exchange but also layer 3 routing or layer 2 switching.
15 [0047] A protocol which is used in cooperation with a routing protocol of a standard network layer to exchange label information between devices on an internet network by label switching is the label distribution protocol (LDP).
[0048] It is to be noted that reference characters 9S and 9D in FIG. 2(a) denote each a terminal such as a personal computer.
20 [0049] In such an MPLS network as shown in FIG. 2(a), a label coordinated with a path leading to a destination terminal is added to an IP packet arriving in the MPLS domain by an edge LSR 11E. In FIG. 2(b), a label A is provided to the IP packet, and the IP packet is transferred to a core LSR 11C. The core LSR 11C looks over only the label and switches the IP packet to an outgoing path corresponding to the label A. When the packet is transferred to the switched outgoing path, a label B is provided newly to the packet. This is called label swapping. In this manner, in the MPLS domain, while label swapping is performed successively, the IP packet is successively transferred toward a target destination terminal 9D in accordance with the label. At the edge LSR 11E at the exit of the MPLS domain, the label
25 is removed and transferred to the MPLS domain.
[0050] Subsequently, a label used in the MPLS is described.
[0051] First, the two following methods are available to realize the MPLS.

- 30 ①Method wherein an existing switch is used
②Method wherein a new label is defined

- [0052] The method according to ①wherein an existing switch is used is a method wherein an existing switch such as an asynchronous transfer mode (ATM) or a frame relay (represented also as FR), and in the ATM, a region of a virtual path identifier (VPI) and a virtual channel identifier (VCI) is used as a label, but in the frame relay, the DLCI is used as a label.
35 [0053] In the method according to ②wherein a new label is defined, a newly defined label called "shim header" is used. The shim header is inserted between the layer 3 and the layer 2, and an outline of it is illustrated in FIG. 3.
[0054] To the shim header, a TTL and a pointer S for a label stack are provided in addition to a label as shown in FIG. 3. The label stack is used for hierarchization of labels, and by using the label stack, it is possible, for example, to 40 hierarchize a plurality of MPLS domains in an MPLS domain to build up a network. In FIG. 3, where S = 1, it is recognized that this is the last label.
[0055] It is to be noted that, in addition to labels for the ATM/FR, also labels for the PPP over SONET and labels for the ether frame (LAN) are defined (refer to the PPP and the LAN shown on the right side in FIG. 3).
[0056] Further, a label has a length of, for example, 4 octets and is indicated as shown in FIG. 4. Such labels can 45 have a hierarchical structure, and the structure is called "label stack". The label and other additional information illustrated in FIG. 4 is sometimes indicated in the MPLS header defined by the MPLS or indicated in the L2 header or the ATM header.
[0057] The S (Bottom of Stack) indicated in FIG. 4 is set to 1 where the entry is the last entry positioned at the lowermost position of the label stack, and it is set to 0 for all of the other entries.
50 [0058] Meanwhile, the TTL (Time-to-Live) shown in FIG. 4 represents a field used to encode a TTL and is decremented one by one before a labeled packet is forwarded. A packet whose TTL becomes zero as a result of the decrementing must not be forwarded any more, and in this instance, either the packet is abandoned as it is, or it is handed to processing of the network layer with the label stack removed therefrom. This packet must not be forwarded also after it is handed to the processing of the network layer.
55 [0059] Subsequently, a label switched path (LSP) is described. The LSP signifies a path from an LSR at the entrance to another LSR at the exit, and may pass intermediate LSRs, an example can be seen also from FIG. 5. A label at each LSR is determined in a coordinated relationship with the LSP. While description is hereinafter given, in the present invention, collection of traffic characteristics is performed for each LSP.

[0060] Further, a method of establishing an LSP is described.

[0061] First, as a method of establishing an LSP, the following two systems have been proposed.

- 5 ①Method wherein the LDP (Label Distribution Protocol) is used
 ②Method wherein the RSVP-tunnel is used

[0062] In the following, a method which uses the RSVP-Tunnel of ②is described.

[0063] The LSP-tunnel is executed expanding the conventional RSVP, and the background of this system arises from that, where the MPLS incorporates the RSVP, a flow can be managed with increased flexibility by unifying them.

10 [0064] The LSP-tunnel defines, for the RSVP, the following five new objects [LABEL REQUEST (used in a path message (PATH message)), LABEL (used in a Resv message), EXPLICIT ROUTE (used in a PATH message), RECORD ROUTE (used in a PATH message or a Resv message) and SESSION ATTRIBUTE (used in a PATH message)].

15 [0065] Here, the LABEL REQUEST object is used in a PATH message and is used to determine a label regarding the path by an intermediate router and a reception router. In order to establish a path (= LSP-tunnel), a source LSR produces a PATH message having a LABEL REQUEST, and further, a node which has transmitted the LABEL REQUEST must make preparations for reception of a Resv message corresponding to the PATH message. If a PATH message which does not include a LABEL REQUEST is received, then a LABEL REQUEST object must not be written into the PATH message.

20 [0066] While a LABEL object is used in a Resv message, in the LSP-tunnel, a label is placed in the Resv message and returned to the source. Then, the Resv message is sent back toward the source along a route of the PATH message. Further, each node which receives the Resv message having the LABEL object uses the label as a label of the output side of the LSP-tunnel, and if the node is not the source of the LABEL REQUEST, then the node executes processing of allocating a new label and placing the new label into the LABEL object of the Resv message. Then, the label transmitted to the upstream is used as an input side label of the LSP-tunnel.

25 [0067] An EXPLICIT ROUTE object (ERO) is used in a PATH message. If this object appears in a PATH message, then the PATH message must be transferred along a path explicitly indicated by the ERO. The ERO is intended to be used for explicit routing. The RECORD ROUTE object is used in a PATH message or a Resv message and allows a transmission node to receive information regarding an actual path of the LSP-tunnel. This is similar to a path vector (path vector) and can be used also for detection of a loop.

30 [0068] A SESSION ATTRIBUTE object is used in a PATH message. This SESSION ATTRIBUTE object can be provided to a PATH message and used for identification and diagnosis of a session. Further, this object has pre-emission (preemption), priority, local protection and so forth.

35 [0069] FIG. 6 illustrates a flow of a message of an LSP-tunnel in which the RSVP is used. The RSVP-tunnel is a Downstream-on-demand system since a label is allocated from the downstream. As can be recognized also from FIG. 6, a first MPLS router (LSR1) in an MPLS domain places a LABEL REQUEST object into a PATH message of the RSVP and transmits the PATH message to the downstream. Thereafter, if the PATH message arrives at a node (LSR4) of the destination, then the destination node (LSR4) places a LABEL object into a Resv message and returns the Resv message to the upstream. The manner of allocation of a label is such as described hereinabove.

40 [0070] In the present invention, a traffic characteristic value is embedded into and transferred together with the RSVP message.

(2-2) Description of the Load Distribution Method in the MPLS

45 [0071] In order to perform load distribution using the MPLS, using the LSP described hereinabove, a plurality of transmission paths [path, i.e., LSP; in the present embodiment, a path and an LSP are used in the same meanings] 12-i (i = 1 to N: N is a natural number) are set between a source MPLS router (a source communication apparatus) 11S and a destination MPLS router (a destination communication apparatus) 11D as shown in FIG. 7. In this instance, the LSP is set using a label of the MPLS. For the setting of the LSP, an LSP setting protocol such as the RSVP-LSP-Tunnel is used. It is to be noted that, in FIG. 7, reference character 11R denotes a repeating MPLS router (an intermediate communication apparatus).

[0072] Further, a portion of a transmission path between routers is called link, and, for example, a transmission path 12-1 in FIG. 7 is composed of M links (link_1 to link_M).

55 [0073] It can also be recognized that the network architecture of the MPLS shown in FIG. 7 may be an IP network which includes a source MPLS router 11S, a destination MPLS router 11D, a plurality of transmission paths 12-i (i = 1 to N: N is a natural number) which can be set between the source MPLS router 11S and the destination MPLS router 11D, and a repeating MPLS router 11R interposed in one of the transmission paths 12-i. It is to be noted that, also in the present embodiment, where the source MPLS router 11S, destination MPLS router 11D and repeating MPLS router

11R are represented without specific distinction among them, they are referred to merely as MPLS routers 11.

[0074] FIG. 8 shows a functional block diagram of an exemplary embodiment of an MPLS router which forms an IP network. As shown in FIG. 8, the MPLS router 11 includes an IP packet transfer section 111, a routing protocol section 112, a path selection section 113, an LSP selection section 114, an LSP setting section 115 and a traffic engineering section 116.

[0075] Here, the IP packet transfer section 111 has a function of performing desired processing such as IP packet transfer processing for an input packet and transferring the packet to another router. The routing protocol section 112 mediates notification of traffic characteristics to another router or collection of traffic characteristics from another router. It is to be noted that the routing protocol section 112 further has a function of rendering, when the MPLS router is connected to the network, the routing protocol operative to discover the other MPLS routers connected to the network. For the discovery, a Hello packet is used usually.

[0076] The path selection section 113 has a function of producing a link state database (Link State Database) 113A for path selection when the topology of the network is discriminated by the routing protocol. The link state database 113A cooperates with the LSP setting section 115 through the LSP selection section 114 to set an LSP with another MPLS router.

[0077] The LSP selection section 114 has a function of cooperating with the link state database 113A of the path selection section 113 to produce an alternate path information database 114A for LSP selection. At this time, a routing table unique to the MPLS which is not used in conventional routing protocols is produced. The alternate path information database 114A cooperates with the traffic engineering section 116 to set an alternative path using an LSP as needed.

[0078] In short, a plurality of LSPs to an MPLS router at the end are set so that load may be balanced between the LSPs.

[0079] The LSP setting section 115 has a function of setting a plurality of LSPs to an MPLS router at the end in response to a result of LSP selection by the LSP selection section 114. A result of the LSP setting by the LSP setting section 115 is transmitted to the IP packet transfer section 111, and a transfer form of an IP packet is set by the IP packet transfer section 111.

[0080] The traffic engineering section 116 has a series of functions regarding a load balance according to the present invention. As shown in FIG. 9, the traffic engineering section 116 includes a load observation section 116A, a traffic characteristic value calculation section 116B, a traffic engineering calculation section 116C and a load adjustment section 116D.

[0081] The load observation section 116A has a function of sending traffic characteristic values such as a utilization and a packet loss number sent thereto from the packet transfer section 111 to the traffic characteristic value calculation section 116B. The traffic characteristic value calculation section 116B has a function of performing smoothing or maximization of the traffic characteristic values and transmitting a resulting value to another MPLS router and another function of collecting traffic characteristic values sent thereto from another MPLS router and sending them to the traffic engineering calculation section 116C.

[0082] The traffic engineering calculation section 116C has a function of performing calculation of an effective load, an effective bandwidth and so forth and performing adjustment of load in response to the bandwidths of the individual LSPs. A result of the calculation (operation) obtained by the traffic engineering calculation section 116C is transmitted to the load adjustment section 116D.

[0083] The load adjustment section 116D has a function of determining a width of a traffic share based on the information from the traffic engineering calculation section 116C and sending a result of the determination to the LSP setting section 115 through the LSP selection section 114. The LSP setting section 115 thereafter sets a traffic share width at the IP packet transfer section 111.

[0084] In particular, the traffic engineering section 116 has functions of a traffic characteristic collection section for collecting traffic characteristics of LSPs (transmission paths) connected to the LSP router itself or the other MPLS routers, a traffic characteristic notification section for notifying the other MPLS routers of the traffic characteristics collected by the traffic characteristic collection section, a load calculation section for calculating load based on the traffic characteristics collected by the traffic characteristic collection section, a decision section for deciding whether or not an LSP should be added or deleted based on the load information determined by the load calculation section, and a load equalization section for equalizing the load determined by the load calculation section among the plurality of LSPs (i is a natural number). This traffic engineering section 116 makes a principal part of the transmission path controlling apparatus of the present invention.

[0085] It is to be noted that principally the load observation section 116A and the traffic characteristic value calculation section 116B exhibit the function of the traffic characteristic collection section 3, the traffic engineering section 116 exhibits the function of the traffic characteristic notification section 4, and the traffic engineering calculation section 116C exhibits the functions of the load calculation section 5, decision section 6 and load equalization section 7.

[0086] Further, in order to allow an MPLS router to have the functions of the traffic characteristic collection section, traffic characteristic notification, load calculation section, decision section and load equalization described above, normally from a medium on which a transmission path controlling program for causing a computer to function as traffic

EP 1 130 849 A2

characteristic collection unit for functioning as the traffic characteristic correction section described above (the traffic characteristic collection unit has a function of collecting traffic characteristics of those of the transmission paths connected to the MPLS router or the other MPLS routers), traffic characteristic notification unit for functioning as the traffic characteristic notification section (the traffic characteristic notification unit has a function of notifying the other MPLS routers of the traffic characteristics collected by the traffic characteristic collection unit), a load calculation unit for functioning as the load calculation section (the load calculation unit has a function of calculating load based on the traffic characteristics collected by the traffic characteristic collection unit), a discrimination unit for functioning as the decision section (the discrimination unit has a function of discriminating whether or not a transmission path should be added or deleted based on the load information determined by the load calculation unit), and a load equalization unit for functioning as the load equalization section (the load equalization unit has a function of equalizing the load determined by the load calculation unit among the plurality of transmission paths) is recorded, the transmission path controlling program is installed into each of the MPLS routers.

[0086] The installation may be performed by setting a recording medium such as a floppy disk or an MO disk to a computer or by using a transmission path controlling program distributed over a network (for example an IP network).

[0087] It is to be noted that, since an actual IP network may be formed like a cobweb using an MPLS router as a node, any MPLS router can serve as a source MPLS router, a destination MPLS router or a repeating MPLS router, and consequently, although the present embodiment is shown as an example wherein the traffic engineering section 116 of each of the source MPLS router 11S, destination MPLS router 11D and repeating MPLS router 11R shown in FIG. 7 has all of the functions of the traffic characteristic collection section, traffic characteristic notification section, load calculation section, decision section and load equalization section described above, it is otherwise possible that the traffic engineering section of the source MPLS router 11S has the functions of the traffic characteristic collection section, traffic characteristic reception section, load calculation section, decision section and load equalization, and the traffic engineering section provided in each of the destination MPLS router 11D and repeating MPLS router 11R other than the source MPLS router 11S has only the functions of the traffic characteristic collection section and traffic characteristic notification section.

[0088] In the following, a load distribution algorithm performed in the traffic engineering section 116 is described in detail.

(2-3) Load Distribution Algorithm Performed in Each MPLS router 11

[0089] First, the MPLS router 11 performs, for each port thereof, traffic collection regarding required traffic items [for example an output packet amount from an output port, an abandoned packet number (NLOSS) at the output port, a logical bandwidth (LBW) of the output port and so forth] as traffic characteristic values after each time t_l [seconds (sec)] (procedure S1).

[0090] Then, the MPLS router 11 calculates a transmission line utilization CUTY at present from the obtained traffic characteristic values in accordance with the following expression (procedure 2).

$$CUTY = \frac{\text{output packet amount from output port}}{\text{logical bandwidth of output port}} \quad (1)$$

[0091] Hence, it follows that the traffic characteristic collection section has means for calculating utilizations of the transmission paths connected to the router based on the collected traffic characteristic values. As a result, in what condition each of the transmission paths is used can be grasped in detail.

[0092] At this time, also working processing of information (the transmission line utilization CUTY) regarding the obtained traffic characteristic values is performed.

[0093] The processing manner is such as follows.

[0094] ① Smoothing processing is performed. Such smoothing processing is performed, as an example, by calculating the following expression each time an observation value is obtained to obtain a value MUTY.

$$MUTY = \alpha \times CUTY + (1 - \alpha) \times MUTY \quad (2)$$

where α is a smoothing coefficient.

[0095] Consequently, it is considered that the traffic characteristic collection section has means for smoothing information (the transmission line utilization CUTY) regarding collected traffic characteristic values. More particularly, it is considered that the traffic characteristic collection section has means for smoothing, after traffic characteristics of

transmission paths connected to the MPLS router are collected, the statistical values in accordance with, for example, the Exponential Moving Average method. As a result, an influence of a sudden traffic variation can be reduced.

[0096] Further, since the transmission path utilization CUTY at present determined from the obtained traffic characteristic values is smoothed as described above, it follows that a smoothing unit may be constructed so as to smooth the utilizations calculated by the utilization calculating unit. Consequently, the utilizations of the individual transmission paths can be grasped accurately while eliminating traffic variation influences.

[0097] ②Maximum value extraction processing is performed. Such maximum value extraction processing is performed by calculating a maximum value of observation values. In other words, the maximum value of the observation values is placed as the work value MUTY. This may be represented in an expression as follows.

10

$$\text{MUTY} = \max(\text{observation values}) \quad (3)$$

[0098] Consequently, it is considered that the traffic characteristic collection section has means for determining a maximum value of collected characteristic values as a representative value of the traffic characteristic values. As a result, transmission path selection control can be performed within a degree of certainty.

[0099] It is to be noted that, as regards the smoothing processing and the maximum value extraction processing described above, both of the smoothing processing and the maximum value selection processing may be performed or only one of the smoothing processing and the maximum value selection processing may be performed.

[0100] ③A total number TLOSS of abandoned packets may be calculated in accordance with the following expression.

[0101] $TLOSS(n) = TLOSS(n-1) + NLOSS$ where $TLOSS(n)$ is the total number of abandoned packets after updated, $TLOSS(n-1)$ is the total number of abandoned packets before updated, and $NLOSS$ is the number of newly abandoned packets.

[0102] Thereafter, the MPLS router 11 executes distribution of the track characteristic values (MUTY, TLOSS and so forth) after each time $t2$ [sec] making use of Flooding of the routing protocol.

[0103] It is to be noted that $TLOSS$ is reset (that is, to 0) after completion of the flooding.

[0104] Also where the maximum value is adopted for MUTY, MUTY is reset (that is, to 0).

[0105] Consequently, it is considered that the traffic characteristic notification section has means for notifying information regarding traffic characteristics utilizing a packet prepared by the routing protocol. More particularly, it is considered that the traffic characteristic notification section has means for distributing, after traffic characteristics of the transmission paths connected to the MPLS router are collected, the characteristic values making use of a packet prepared by the routing protocol such as the OSPF or the IBGP to notify another MPLS router of the information. As a result, an existing protocol can be utilized, and consequently, it is not necessary to newly prepare a routing protocol for exclusive use, which contributes to reduction of the cost and so forth.

[0106] Also it is considered that the traffic characteristic notification section has means for notifying information regarding traffic characteristics after each fixed period $t2$ [sec]. As a result, the arithmetic load applied to an arithmetic unit such as a CPU which executes arithmetic operation and so forth can be reduced.

[0107] Further, in the present embodiment, it is considered that the traffic characteristic notification section has means for notifying information regarding traffic characteristics with a message of the RSVP expanded. More particularly, it is considered that the traffic characteristic notification section has means for distributing, after traffic characteristics of the transmission paths connected to the MPLS router are collected, the characteristic values with a message of the RSVP expanded to notify another MPLS router of the information. As a result, an existing protocol can be utilized, and consequently, it is not necessary to newly prepare a routing protocol for exclusive use, which contributes to reduction of the cost and so forth.

[0108] Further, it is considered that the traffic characteristic collection section has means for collecting a unit of a flux of flows for collection of traffic characteristics in a unit of a label used in the MPLS. As a result, collection of fine traffic characteristic becomes possible, which contributes to augmentation of the collection accuracy.

[0109] It is to be noted that similar effects can be achieved also by an alternative construction wherein calculation of the transmission path utilizations at present is performed after some or all of the smoothing processing, the maximum value extraction processing and the total abandoned packet number calculation processing described above are performed for the collected traffic characteristic values.

[0110] On the other hand, also it is possible to employ a further construction wherein the calculation processing of the transmission path utilizations is assigned to the MPLS router 11S which serves as a start point of load distribution while the repeating MPLS routers and the destination router perform only the other processing (smoothing processing, maximum value extraction processing, total abandoned packet number calculation processing and traffic characteristic value notification processing) than the transmission path utilization calculation processing.

[0111] A load distribution algorithm performed by a traffic engineering section 116' of each of the repeating MPLS

routers 11R' and the destination router 11D'(the routers mentioned are hereinafter denoted generally by 11') having the construction just described is described in detail below.

[0112] First, similarly to the traffic engineering section 116 described above, also the traffic engineering section 116' of the MPLS router 11' performs, for each port thereof, traffic collection regarding required traffic items [for example an output packet amount (L) from an output port, an abandoned packet number (NLOSS) at the output port, a logical bandwidth (LBW) of the output port and so forth] as traffic characteristic values after each time t1 [seconds (sec)].

[0113] Then, the traffic engineering section 116' performs smoothing processing of the obtained output packet amounts. Such smoothing processing is performed by obtaining a value ML by calculating the following expression each time an observation value is obtained:

10

$$ML = \alpha \times L + (1 - \alpha) \times ML' \quad (5)$$

where α is a smoothing coefficient, and ML' is ML calculated in the preceding cycle.

15

[0114] Then, maximization extraction processing is performed. This processing is performed by placing a maximum value of the observation values as a work value ML . It is to be noted that the maximization extraction processing can be omitted.

[0115] Further, similarly to the traffic engineering section 116 described above, the traffic engineering section 116' calculates a total number TLOSS of abandoned packets in accordance with the expression (4).

20

[0116] Thereafter, the traffic engineering section 116' executes distribution of the traffic characteristic values (ML , TLOSS and so forth) after each time t2 [sec] making use of Flooding of the routing protocol.

[0117] It is to be noted that, similarly to the traffic engineering section 116, the traffic engineering section 116' resets TLOSS (that is, to 0) after each completion of the flooding and, where the maximum value is adopted as ML , the traffic engineering section 116' resets ML (that is, to 0) as well.

25

[0118] Consequently, similarly to the MPLS router 11 described hereinabove, also the MPLS router 11' of the present modification can achieve those effects provided by the smoothing processing, the maximum value extraction processing, the total abandoned packet number calculation processing and the traffic characteristic value notification processing. Further, since the smoothing processing of information regarding the traffic characteristics is assigned to the MPLS router 11S which serves as a start point of load distribution, the time required to work the traffic characteristics by the MPLS router 11' can be reduced from that by the MPLS router 11 described above and the MPLS router 11' can be simplified in construction, which contributes to reduction of the cost and so forth.

30

(2-4) Load Distribution Algorithm Performed by the MPLS Router 11S Serving as a Start Point of Load Distribution

35

[0119] First, flooding information of which the MPLS router 11S which is determined as a start point of load distribution is notified includes the average link utilization (or maximum value of observation values) MUTY (or average output packet amount ML), abandoned packet number TLOSS, output port logical bandwidth LBW and so forth as described hereinabove.

40

[0120] Here, where the MPLS router 11S receives the average output packet amount ML , the link utilization MUTY of each router is calculated using an output port logical bandwidth LBW and so forth received together. Whether the other intermediate routers have an average link utilization calculation function and issues a notification of the calculated average link utilization MUTY, the link utilization calculation function need not to be provided for the MPLS router 11S side.

45

[0121] The MPLS router 11S which serves as the start point of load variation performs path state checking, calculation of an effective load and load adjustment.

[0122] As regards the path state checking, congestion state checking of the individual routes from the MPLS router (start point router) 11S to the MPLS router (end point router) 11D is performed in the following manner after each fixed period t3 [sec].

50

[0123] On the premise that a plurality of paths are laid between the MPLS router 11S which servers as the start point of load distribution control and the MPLS router 11D which is determined as an end point of the load distribution control and each path extends through a plurality of repeating MPLS routers 11R, the congestion state checking for each route discriminates, from utilizations of each path, collected from each of the repeating MPLS routers 11R of the path, based on an average utilization of the transmission paths (links) connected to the repeating MPLS router 11R [this congestion state checking manner is referred to as congestion state checking manner where utilizations of links j (link j) of a path i are averaged].

[0124] This can be represented in such an expression as given below.

$$\rho_{\text{path } i} = \text{Average}(\text{MUTY}(\text{link } j, \text{path } i)) \quad (6)$$

- 5 [0125] Consequently, in this instance, it is considered that the traffic characteristic collection section provided in the source MPLS router 11S has means for discriminating a utilization of each transmission path connected to the source MPLS router based on an average utilization of the transmission paths connected to repeating MPLS routers collected from the repeating MPLS routers. As a result, this can contribute to augmentation of the efficiency in use of the network.
- 10 [0126] Further, on a premise similar to that in the congestion state checking manner where utilizations of links j of a path i are averaged described above, the congestion state checking for each route may be such congestion state checking that it discriminates, from utilizations of each path, collected from each of the repeating MPLS routers, based on a maximum utilization of the transmission paths (links) connected to the repeating MPLS router (this congestion state checking manner is referred to as congestion state checking manner where a maximum value of utilizations of all links of a path i is adopted).

15 [0127] This can be represented in the following expression.

$$\text{path } i = \text{Max}(\text{MUTY}(\text{link } j, \text{path } i)) \quad (7)$$

- 20 [0128] Consequently, in this instance, it is considered that the traffic characteristic collection section provided in a source MPLS router has means for discriminating a utilization of each transmission path connected to the source MPLS router based on a maximum utilization of transmission paths connected to repeating MPLS routers collected from the repeating MPLS routers. As a result, transmission path selection control can be performed within a degree of certainty.
- 25 [0129] It is to be noted that the sum total TLOSS path i of abandoned packets of the path i may be determined in accordance with the following expression.

$$\text{TLOSS path } i = \sum \text{TLOSS link } j \quad (8)$$

- 30 [0130] It is to be noted that i above is the number of the path (LSP) and j is the link number.
- [0131] Also it is considered that the traffic characteristic collection section provided in a source MPLS router has means for discriminating after each fixed period $t3$ [sec] whether or not each transmission path is in congestion. By this, the arithmetic load applied to an arithmetic unit such as a CPU which executes arithmetic operation and so forth can be reduced.
- 35 [0132] Calculation of the effective load (Effective load) $\rho_{\text{effective path } i}$ may be performed in the following manner.
- [0133] Here, the effective load is defined as an effective utilization calculated from a utilization of a link and a packet loss number at the link. Originally, an actual load applied to the link should be measured. However, the inside of a router actually has a multi-stage switch construction and it is difficult to directly measure the load. Accordingly, in the present embodiment, a load is estimated using only a utilization of a link and a packet loss number, and the estimated load is referred to as effective load.
- 40 [0134] The effective load is described with reference to FIG. 10. In particular, as can be seen apparently from a characteristic of FIG. 10, if no packet loss is involved, then the load $\rho_{\text{path } i}$ and the effective load at the link are equal to each other. However, if packet loss occurs, then the load $\rho_{\text{path } i}$ is corrected with a predetermined function $f(\text{TLOSS path } i)$ so that a rather high load may be obtained, and a resulting value is used as the effective load. It is to be noted that an upper limit ρ_{ceiling} is provided to the effective load. This can be represented in the following expression.

$$\rho_{\text{effective path } i} = \rho_{\text{path } i} \times f(\text{TLOSS path } i) \quad (9)$$

- 45 $\rho_{\text{effective path } i} = \text{Min}(\rho_{\text{effective path } i}, \rho_{\text{ceiling}})$ where $f(\text{TLOSS path } i)$ is a function regarding TLOSS path i , and ρ_{ceiling} is an upper limit load, that is, an upper limit value to the effective load $\rho_{\text{effective path } i}$.
- 50 [0135] This function is exhibited in the load calculation section provided in a source MPLS router for calculating an effective load taking a number of abandoned packets generated in a repeating MPLS router into consideration. Consequently, in a situation that a plurality of paths are laid between an MPLS router which is determined as a start point of load distribution control and another MPLS router which is determined as an end point of the load distribution control and each path extends through a plurality of repeating MPLS routers, when the MPLS router which serves as the start point of the load distribution control calculates a load based on traffic characteristics notified of, it can calculate, taking the number TLOSS path i of abandoned packets generated in each repeating MPLS router into consideration, the effective load $\rho_{\text{effective path } i}$ in accordance with the following equation.

effective load to ith path = load to ith path × f(abandoned packet number)

It is to be noted that f (abandoned packet number) is a function wherein the abandoned packet number is a variable.

- 5 [0136] As a result, even if a load is not directly measured, an effective load can be estimated by a simple method.
- [0137] Further, in another embodiment, it is considered that the effective load calculation unit is constructed such that, when it calculates an effective load, it sets an upper limit to the effective load. Consequently, load estimation of an excessive amount can be avoided, and as a result, appropriate transmission path selection control can be performed.
- [0138] The load adjustment is performed in the following manner.
- 10 [0139] The load adjustment is executed after each time t4 [sec], and if the effective load to each path (LSP) can be calculated in such a manner as described above, then a utilization at all paths (LSP) is determined subsequently. In the present embodiment, all paths are regarded as a single virtual pipe, and a load to the pipe is calculated.
- [0140] In particular, in this instance, it is considered that the decision section provided in a source MPLS router has means for calculating effective bandwidths from effective loads obtained by the effective load calculation unit and regarding all of the transmission paths as one virtual pipe and then determining a utilization of the virtual pipe. In this instance, the utilization may be determined in accordance with the following expression.

utilization ρ effective all when all paths are

- 20 regarded as one virtual pipe = $\Sigma (\rho \text{ effective path } i \times LBW$

$$\text{path } i)/\Sigma LBW \text{ path } i \quad (11)$$

- 25 where ρ effective path is the effective load of the path i , and $LBW \text{ path } i$ is the logical bandwidth of the path i .
- [0141] Consequently, it can be prevented to unnecessarily add a new transmission path, and this can contribute also to effective utilization of the network.
- [0142] Subsequently, effectiveness of multi-paths is verified.
- 30 [0143] In particular, the utilization obtained from the expression (11) is compared with a set reference value ρ offer1 for path addition, and when the obtained value of the utilization is higher than the reference value for path addition, a new transmission path is added. Further, the utilization obtained from the expression (11) is compared with a set reference value for path deletion similarly, and if the obtained value of the utilization is lower than a reference value ρ offer2 for path deletion, then a path whose load is low is deleted. In particular, if the calculated load described above is high, then a new path is added, but if it is low, an existing path is deleted.
- 35 [0144] This can be logically represented in the following manner.
- [0145] If ρ effective all $\geq \rho$ offer1 then

```

if timer1 was expired then (the left signifies "if the time has expired already")
search new path and add new path (the left signifies "to the alternative path search processing")
endif
elseif  $\rho$  effective all <  $\rho$  offer2 then
if timer2 was expired then (the left signifies "the timer has expired already")
delete path (the left signifies "to the alternative path deletion processing")
endif
endif

```

- 45 [0146] Consequently, it is considered that the decision section provided in a source MPLS router has means for comparing a pipe utilization obtained by the pipe utilization calculation unit with a set reference value to perform discrimination of whether or not a transmission path should be added or deleted in response to a result of the comparison, and adds a new transmission path if the pipe utilization obtained by the pipe utilization calculation unit is higher than a reference value for path addition, but deletes a transmission path whose load is low if the pipe utilization obtained by the pipe utilization calculation unit is lower than a reference value for path deletion. As a result, up-to-date optimum path selection can be performed.
- [0147] It is to be noted that, as a modification to the verification procedure of the validity of multi-paths described above, it is otherwise possible to calculate, in addition to the pipe utilization (total utilization of all paths) obtained using the expression (11), a rate of change of the pipe utilization per time and perform decision of addition or deletion of a transmission path based on the rate of change of the pipe utilization. A procedure in this instance is described in detail below.

[0148] First, a difference value $\Delta U(t)$ between two pipe utilizations obtained after a very small interval t_4 is calculated as the rate of change of the pipe utilization per time in accordance with the following expression. =

$$5 \quad \Delta U(t) = \{ p_{\text{effective all}}(t-1) - \\ p_{\text{effective all}}(t) \} / t_4 \quad \dots (12)$$

10 where $p_{\text{effective all}}(t)$ is the total utilization of all paths at time t , and $p_{\text{effective all}}(t-1)$ is the total utilization at a measurement time $t-1$ prior by one cycle to time t .

[0149] It is to be noted that it is otherwise possible to represent a change of the pipe utilization with respect to time obtained in accordance with the expression (11) as a function of the time and calculate a differential value of the pipe utilization at a certain point of time and then use the differential value as the rate of change of the pipe utilization per time. This allows a more accurate rate of change of the pipe utilization to be grasped.

[0150] Then, the utilization obtained in accordance with the expression (11) and a reference value p_{offer1} for path addition set therefore are compared with each other, and the rate of change obtained in accordance with the expression (12) and a reference value p_{offerd1} for path addition set therefore are compared with each other. Then, when the values of the obtained utilization and rate of change both exceed the corresponding reference values for path addition, 20 a new transmission path is added. Further, the utilization obtained in accordance with the expression (11) and a reference value p_{offer2} for path deletion set therefore are compared with each other, and the rate of change obtained in accordance with the expression (12) and a reference value p_{offerd2} for path deletion set therefore are compared with each other. Then, when the value of the obtained utilization or rate of change is smaller than the corresponding reference value for path deletion, a path which has a comparatively low load is deleted. In other words, also in this instance, if 25 the load calculated as above is high, then a new path is added, but if the load is low, then an existing path is deleted.

[0151] This can be logically represented in the following manner.

[0152] If ($p_{\text{effective all}} \geq p_{\text{offer1}}$ or $\Delta U(t) \geq p_{\text{offerd1}}$) then

```
30   if timer1 was expired then (the left signifies "if the timer has expired already")
      search new path and add new path (the left signifies "to the alternative path search processing")
      endif
      elseif ( $p_{\text{effective all}} < p_{\text{offer2}}$  and  $\Delta U(t) < p_{\text{offerd2}}$ ) then
        if timer2 was expired then (the left signifies "the timer has expired already")
          delete path (the left signifies "to the alternative path deletion processing")
        endif
      endif
```

[0153] Through the construction described above, since decision of addition or deletion of a path is performed with a change in time of the pipe utilization taken into consideration, addition or deletion of a path can be performed more effectively than where such decision is based only on the pipe utilization.

[0154] It is to be noted that, while, in the construction described above, logical ANDing (and) of results of comparison is used for discrimination of addition of a path and logical ORing (or) of results of comparison is used for discrimination of deletion of a path, the criteria for addition/deletion of a path are not limited to the combination given above, but those criteria (logical ANDing and logical ORing) may be used in a different combination in accordance with a configuration 45 or a state of the network, or some other criteria may be used. This allows an integrated amount of the pipe utilization to be grasped appropriately in accordance with the configuration or the state of the network.

[0155] Meanwhile, as another modification to the verification procedure of the validity of multi-paths, also it is possible to calculate, in addition to the pipe utilization (total utilization of all paths) obtained using the expression (10), an integral amount of the pipe utilization in a fixed time and perform discrimination of addition or deletion of a transmission path 50 based on the integral amounts of the pipe utilization.

[0156] In this instance, as an integral amount of the pipe utilization in a fixed time, an integral value $\int MU(t)$ of the pipe utilization within a fixed time t_4 may be calculated in accordance with the following expression.

[Expression 1]

[0157]

5

$$\int MU(t) = \int_{t_1}^{t+t_4} MU(t)dt \quad \dots (13)$$

- 10 [0158] In other words, $\int MU(t)$ is an integral value of the pipe utilization MU within the fixed time from t to t_4 .
 [0159] Also it is possible to calculate a number of times by which the pipe utilization exceeds a certain reference value within the fixed period of time and use the number of times as the integrated amount of the pipe utilization within the fixed period of time. This allows the integrated amount of the pipe utilization to be grasped by a simpler construction.
 15 [0160] Then, the utilization obtained in accordance with the expression (11) and a reference value p_{offer1} for path addition set therefore are compared with each other, and the integrated amount obtained in accordance with the expression (13) and a reference value $p_{offer1l}$ for path addition set therefore are compared with each other. Then, a new transmission path is added based on results of the comparison. Further, the utilization obtained in accordance with the expression (11) and a reference value p_{offer2} for path deletion set therefore are compared with each other, and the integrated amount obtained in accordance with the expression (13) and a reference value $p_{offer2l}$ for path deletion set therefore are compared with each other. Then, a path which has a comparatively low load is deleted based on results of the comparison.
 20 [0161] Also in this instance, upon decision of addition or deletion of a path, various criteria such as logical ORing and logical ANDing of the results of comparison may be used in a suitable combination in accordance with the configuration or the state of the network, and this allows the integrated amount of the pipe utilization to be grasped appropriately. Further, where discrimination is performed based only on the integrated amount without using the utilization, decision of addition or deletion of a path can be performed with a simpler construction.
 25 [0162] From the foregoing construction, it follows that the decision section provided in the source router has a calculation unit for calculating an effective bandwidth from the effective load obtained by the effective load calculation unit and for determining, regarding all of the transmission paths as one virtual pipe, information regarding a utilization of 30 the virtual pipe (a pipe utilization, a rate of change of the pipe utilization, an integrated amount of the pipe utilization) and is constructed so as to perform decision of whether or not a transmission path should be added or deleted based on the information regarding the utilization of the virtual pipe (at least one of the pipe utilization, the rate of change of the pipe utilization, and the integrated amount of the pipe utilization). Consequently, it is possible to perform addition and deletion of a transmission path appropriately taking the loads of the transmission paths at present into consideration, and distribution of the load can be anticipated efficiently and selection of an appropriate path can be performed 35 in accordance with the situation of the network or the performance of the hardware.
 [0163] Further, similarly to the foregoing construction, it follows that the decision section provided in the source router is constructed so as to perform comparison of at least one of a set of the pipe utilization obtained by the calculation unit and a first reference value (p_{offer1} , p_{offer2}), another set of the rate of change obtained by the calculation unit 40 and a second reference value ($p_{offerd1}$, $p_{offerd2}$) and a further set of the integrated value obtained by the calculation unit and a third reference value ($p_{offer1l}$, $p_{offer2l}$), and perform decision of whether or not a transmission path should be added or deleted in response to a result of the comparison. More particularly, when at least one of the utilization, the rate of change and the integrated amount exceeds a corresponding reference value for path addition, it is decided 45 that a new transmission path should be added, but when at least one of the utilization, the rate of change and the integrated amount is lower than a corresponding reference value for path deletion, it is decided that a transmission path whose load is comparatively low should be deleted. Consequently, up-to-date optimum path selection can be performed in accordance with the situation of the network or the performance of the hardware.
 [0164] By the way, if it is decided upon verification of the validity of multi-paths described above that a transmission 50 path should be deleted, then a path of a candidate for deletion is selected based on several conditions, and the utilization of all paths (loads to the other transmission paths) when the path of the candidate for deletion is deleted is estimated and compared with a predetermined reference value. Then, a path to be deleted actually is determined based on a result of the comparison.
 [0165] More particularly, deletion of multi-paths is performed in accordance with the following algorithm.

55

- ①That one of all paths which has been laid first is excluded from the object of deletion.
- ②That one of all paths which has the smallest usable bandwidth is selected.
- ③If a plurality of paths satisfy the condition of ②(have the smallest usable bandwidth), then that one of the paths which has the lowest effective load is selected.

EP 1 130 849 A2

④ If a plurality of paths satisfy the conditions of ② and ③ (have the smallest usable bandwidth and the lowest effective load), then that one of the paths which has the greatest hop number is selected.

⑤ If a plurality of paths satisfy the conditions of ②, ③ and ④ (have the smallest usable bandwidth, the lowest effective load and the greatest hop number), then that one of the paths which has been added last is selected.

5 ⑥ The path selected through ①- ⑤ is determined as a candidate for deletion, and the utilization of the entire multi-paths when the path is virtually deleted is calculated.

⑦ The utilization of the entire multi-paths after the deletion of the deletion candidate path calculated in ⑥ is compared with a congestion discrimination threshold value set therefore. If the utilization of the entire multi-paths after the deletion is equal to or higher than the threshold value, then a new deletion candidate path is selected through 10 ①- ⑤ again from among the paths other than the deletion candidate path.

⑧ If the comparison in ⑦ reveals that the utilization of the entire multi-paths after the deletion is lower than the congestion discrimination threshold value, then the deletion candidate path is deleted actually.

15 [0166] By performing deletion of a path in accordance with such an algorithm as described above, it follows that the decision section provided in the source router includes means for selecting a candidate of a transmission path as an object of deletion based on the load information determined by the load calculation unit, and for estimating the load to the other transmission paths when the transmission path of the candidate of deletion selected by the deletion candidate selection unit is deleted, and for comparing the load of the other transmission paths estimated by the load estimation unit with a predetermined reference value and determining whether or not the transmission path of the candidate of deletion should be deleted in response to a result of the comparison.

20 [0167] It is to be noted that, when the result of comparison between the information (utilization, rate of change, integrated amount) regarding the pipe utilization (total utilization of all paths) described hereinabove and the reference value is linked to the determination that a transmission path should be deleted (for example, when one of the different kinds of information regarding the pipe utilization is lower than the corresponding reference value), the path deletion algorithm described above is executed using the result of comparison as a trigger so that deletion of a path is performed 25 actually.

[0168] From the construction described above, it follows that the decision section provided in the source router further includes a trigger unit for triggering the selection unit, estimation unit and determination unit to operate based on the information regarding the pipe utilization obtained by the calculation unit.

30 [0169] Consequently, a transmission path can be deleted in an up-to-date manner in response to a situation of the entire network.

[0170] After the validity of multi-paths is verified in this manner, an effective bandwidth (EBW) to be moved is calculated.

35 [0171] In particular, when to equalize load between a plurality of transmission paths added, in order to distribute the load in proportion to the logical bandwidths of the paths, the effective bandwidth is calculated in accordance with an expression (14) to perform load distribution.

$$\text{effective bandwidth to be moved of } i\text{th path} =$$

40
$$(\text{average value of effective bandwidths of all paths}$$

$$- \text{effective bandwidth of } i\text{th path})$$

45
$$\times \text{logical bandwidth of } i\text{th path} \quad (14)$$

50 [0172] As a result, balancing of the load can be performed between different paths, and a concept diagram for illustrating an effective bandwidth is illustrated in FIG. 11. Referring to FIG. 11, ΔEBW_i indicates an effective bandwidth to be moved, and this effective bandwidth is obtained by multiplying an effective load by the logical bandwidth LBW. As a result, ΔEBW can be calculated using an expression (15).

$$\Delta EBW \text{ path } i = (\rho_{\text{effective all}} - \rho_{\text{effective path}} \times LBW \text{ path } i \quad (15)$$

55 [0173] Consequently, it is considered that the load equalization provided in a source MPLS router has more effective bandwidth calculation means for determining an effective bandwidth to be moved at each of transmission paths from

an average effective bandwidth determined for all transmission paths based on effective bandwidths obtained from effective loads obtained taking numbers of abandoned packets generated in repeating MPLS routers into consideration. Consequently, equalization of load with bandwidths of transmission paths taken into consideration can be performed.

[0174] Then, after the effective bandwidth ΔEBW is determined in such a manner as described above, adjustment of the traffic share is performed finally.

[0175] In particular, IP packets flowing in are distributed to a plurality of paths in proportion to effective bandwidths to be moved calculated in accordance with the expression (14). When the IP packets flowing in are to be distributed to a plurality of paths, calculation by a Hash function is performed based on IP addresses and the IP traffic flow is distributed at random based on a result of the calculation.

[0176] The foregoing is described in more detail in regard to a case wherein a CRC (cyclic redundancy check) calculation is used as the Hash function.

amount Δts path i of traffic to be moved of ith path

$$= \text{effective bandwidth to be moved of ith path} \times Gr \times$$

$$TS/(\text{total of (effective load to ith path} \times \text{logical}$$

$$\text{bandwidth of ith path}) \text{ for all path})$$

$$= \Delta EBW \text{ path } i \times Gr \times TS/\Sigma (\rho \text{ effective path } i \times LBW$$

$$\text{path } i) \quad (16)$$

where Gr is the load adjustment coefficient ($1/p$ ceiling $\leq Gr \leq 1/100$), TS is the width of a result of calculation of CRC arithmetic (in the case of CRC 16, the width of TS is 0-65535).

Gr : load adjustment coefficient

TS : width of traffic share = range of result of CRC 16 = 65535

[0177] When load adjustment is performed in accordance with the expression above, the load adjustment distributions before and after the load adjustment are such as illustrated in FIG. 12.

[0178] In the present embodiment, the track share which has a significance of load adjustment is a parameter used to perform load distribution, and the sum total of traffic shares is a total value of actual traffic through all paths. The width of a traffic share is an integer from 0 to 65,535 where the CRC (cyclic redundancy check) 16 is used. When an IP packet arrives at a router, the load adjustment apparatus can perform CRC calculation based on an IP address (host address, destination address) of the IP packet, and as a result, a certain value between 0 and 65,535 is allocated to the IP flow. In short, the present load adjustment mechanism has a function of degrading an IP address to a width of a certain integer. In other words, IP flows arriving at a router are accommodated at random into individual paths.

[0179] Consequently, it is considered that the load equalization section provided in a source MPLS router is constructed such that it has a unit for distributing packets flowing into the source MPLS router to a plurality of transmission paths in proportion to effective bandwidths to be moved determined by the move effective bandwidth calculation unit and further that it thereupon performs arithmetic operation using a Hash function based on addresses and distributes traffic flows at random based on a result of the arithmetic operation. As a result, traffic flows can be distributed precisely by simple calculation.

[0180] It is to be noted that, as a simpler and higher-speed Hash function, replacement of positions and logical operation may be performed for bit values of an input value to produce a random integer value.

[0181] For example, as illustrated in FIG. 13, a destination IP address of 32 bits included in an IP packet is first divided into four areas of 8 bits. Then, the values of two adjacent 8-bits are replaced alternately, and the upper 16 bits and the lower 16 bits are logically exclusively ORed. Finally, the value of 16 bits produced in this manner is converted into an integer, and the integer is used as a hash value (output value of the Hash function).

[0182] It is to be noted that the replacement of the positions and the logical operation for the bit values of the input value are not limited to those described above, and various methods can be used in an arbitrary combination for them.

[0183] From the foregoing, it follows that the load equalization section provided in the source router is constructed so as to perform at least one of replacement of positions and logical operation for bit values of an input value as the calculation by the Hash function to produce a random integer value. This allows simplification of the hash calculation and also allows distribution of traffic flows to be performed at a high speed.

EP 1 130 849 A2

- [0184] It is to be noted that, in order to augment the efficiency of the hash calculation described above, as an input value for the Hash function, a control value (for example, a protocol ID, a port number or the like) other than the address included in each packet or the like may be used in addition to the address (host address, destination address) included in the packet. In particular, the load equalization section provided in the source router is constructed so as to additionally use, as an input value for the Hash function, a control value other than the address included in each packet. This increases the disturbance degree of the output value by the hash calculation and allows efficient and effective distribution of traffic flows.
- [0185] In the following, an example of load distribution of the present algorithm is described.
- 10 (A) Calculation Example 1
- [0186] If it is assumed that the logical bandwidths LBW [bit/s] of the paths 1, 2 and 3 are 10M, 8M and 2M, respectively, and the effective loads $p_{\text{effective}}$ are 0.5, 0.2 and 0.3, respectively, then the actual traffics [bit/s] of the paths 1, 2 and 3 are $10M \times 0.5 = 5M$, $8M \times 0.2 = 1.6M$, and $2M \times 0.3 = 0.6M$, respectively.
- 15 [0187] Further, if all paths in this instance are regarded as one virtual pipe, then an average utilization $Pave_{\text{effective}}$ is determined in the following manner.

$$Pave_{\text{effective}} = (5M + 1.6M + 0.6M)/(10M + 8M + 2M) = 0.36$$

- 20 [0188] Then, if the effective bandwidth (EBW) [bit/s] to be moved is calculated for the paths 1, 2 and 3, then the following results are obtained.
- [0189] In particular, the effective bandwidth $\Delta EBW_{\text{PATH1}}$ to be moved of the path 1 is $(0.36 - 0.5) \times 10M = -1.4M$; the effective bandwidth $\Delta EBW_{\text{PATH2}}$ to be moved of the path 2 is $(0.36 - 0.2) \times 8M = +1.28M$; and the effective bandwidth $\Delta EBW_{\text{PATH3}}$ to be moved of the path 3 is $(0.36 - 0.3) \times 2M = +0.12M$. Then, if the sum total of the effective bandwidths to be moved of all of the paths is calculated, then the result is $-1.4M + 1.28M + 0.12M = 0$.
- [0190] Adjustment of the traffic share is performed based on the calculation result.
- [0191] In particular, for the path 1, $5M - 1.4M = 3.6M$; for the path 2, $1.6M + 1.28M = 2.88M$; and for the path 3, $0.6M + 0.12M = 0.72M$. Consequently, for the paths 1, 2 and 3, the traffic share should be changed at the ratio of $3.6 \times Gr : 2.88 \times Gr : 1.2 \times Gr$.

(B) Calculation Example 2

- 35 [0192] If it is assumed that the logical bandwidths LBW [bit/s] of the paths 1, 2, 3 and 4 are 150M, 45M, 150M and 600M, respectively, and further the effective loads $p_{\text{effective}}$ are 1.2, 1.5, 1.1 and 0.8, respectively, then the actual traffics [bit/s] of the paths 1, 2, 3 and 4 are $150M \times 1.2 = 180M$, $45M \times 1.5 = 67.5M$, $150M \times 1.1 = 165M$, and $600M \times 0.8 = 480M$, respectively.
- [0193] Further, if all paths are regarded as one virtual pipe, then an average utilization $Pave_{\text{effective}}$ is determined in the following manner.

40

$$Pave_{\text{effective}} = (180M + 67.5M + 165M + 480M)/(150M + 45M + 150M + 600M) = 892.5/945 = 0.94$$

- 45 [0194] Then, if the effective bandwidth (EBW) to be moved is calculated for all of the paths 1 to 4, then the following results are obtained.
- [0195] In particular, the effective bandwidth $\Delta EBW_{\text{PATH1}}$ to be moved of the path 1 is $(0.94 - 1.2) \times 150M = -39M$; the effective bandwidth $\Delta EBW_{\text{PATH2}}$ to be moved of the path 2 is $(0.94 - 1.5) \times 45M = -25.2M$; the effective bandwidth $\Delta EBW_{\text{PATH3}}$ to be moved of the path 3 is $(0.94 - 1.1) \times 150M = -24M$, and the effective bandwidth $\Delta EBW_{\text{PATH4}}$ to be moved of the path 4 is $(0.94 - 0.8) \times 600M = 84M$.
- 50 [0196] Adjustment of the traffic share is performed based on the calculation result.
- [0197] In particular, for the path 1, $180M - 39M = 141M$; for the path 2, $67.5M - 25.2M = 42.3M$; for the path 3, $165M - 24M = 141M$; and for the path 4, $480M + 84M = 564M$. Consequently, for the paths 1, 2, 3 and 4, the traffic share should be changed at the ratio of $141 \times Gr : 42.3 \times Gr : 141 \times Gr : 564 \times Gr$.

55 (2-5) Description of Operation

- [0198] With such a construction as described above, if the MPLS router 11 is first connected to the network, then the routing protocol operates and performs discovery of the other MPLS routers connected to the network. For such

discovery, usually a hello packet is used. If the topology of the network is cleared up by the routing protocol, then a link state database 113A is produced by the path selection section 113. LSPs are set between the MPLS router 11 and the other MPLS routers by the LSP setting section 115 through the LSP selection section 114 using the link state database 113A. Further, the LSP selection section 114 cooperates with the link state database 113A of the path selection section 113 to produce an alternate path information database 114A and thus produces a routing table unique to the MPLS and not used by any other conventional routing protocol. The alternate path information database 114A cooperates with the traffic engineering section 116 to set an alternative path using the LSPs as needed. In short, by setting a plurality of LSPs between the MPLS router 11 and an MPLS router at the end point, balancing of load can be performed between the LSPs.

5 [0199] It is to be noted that, while the series of functions regarding balancing of load are functions of the traffic engineering section 116, to the load observation section 116A of the traffic engineering section 116 having the construction just described, track characteristic values such as a utilization and a packet loss number are sent from the IP packet transfer section 111. The load observation section 116A sends the values to the traffic characteristic value calculation section 116B. The traffic characteristic value calculation section 116B performs smoothing or maximization 10 of the traffic characteristic values and transmits a resulting value to the other MPLS routers. On the other hand, traffic characteristic values sent to the MPLS router 11 from the other MPLS routers are collected by the traffic characteristic value calculation section 116B and sent to the traffic engineering calculation section 116C. The traffic engineering calculation section 116C performs calculation of effective loads and effective bandwidths and performs load adjustment 15 in accordance with the bandwidths of the individual LSPs.

20 [0200] A result of the load adjustment is transmitted to the load adjustment section 116D, by which a width of the track share is determined. Then, a result of the determination is set by the IP packet transfer section 111 through the LSP setting section 115.

25 [0201] In this manner, each of the routers 11R and 11D other than the source MPLS router 11S collects traffic characteristics of transmission paths connected to the other routers and notifies the source router of the collected track characteristics. Meanwhile, the source MPLS router 11S collects traffic characteristics of transmission paths connected to the source router and calculates effective loads based on one or both of the thus collected traffic characteristics and traffic characteristics obtained from the other routers. Then, based on the effective load information, the source MPLS router 11S performs discrimination of whether or not a transmission path should be added or deleted and further performs equalization of the obtained effective load between a plurality of transmission paths. Consequently, each router 30 has a traffic characteristic observation function and a traffic characteristic value notification function, and a router (source router) which serves as a start point of load distribution can dynamically perform bandwidth management of a plurality of paths to another router (destination router) which serves as an end point. As a result, load distribution in accordance with free bandwidths can be performed. In particular, since a plurality of routes (transmission paths) can be set from a router (source router) which serves as a start point to another router (destination router) which serves 35 as an end point from among routers in an IP network and distribution of load can be performed among the set routes, traffic engineering can be achieved within the network such as the Internet irrespective of the network topology and without depending upon the type of the transmission paths.

(2-6) Examples of Experiment

40 [0202] Evaluation of the load distribution performance by the load distribution apparatus of the present invention was performed using a network construction for experiment shown in FIG. 14. It is to be noted that, in FIG. 14, LSR 120R-1 to 120R-5 denote each an LSR which includes the load distribution apparatus of the present invention, and a load generating apparatus 121 is connected to the LSR 120R-1 which serves as a source LSR while a load receiving apparatus 122 is connected to the LSR 120R-5 which serves as a load destination LSR. A plurality of links link1 to link5 are laid between the LSRs 120R-1 to 120R-5 such that they form a plurality of transmission paths 123-1 and 123-2. In other words, also the network construction for experiment shown in FIG. 14 can be considered as an example of the network construction of the MPLS shown in FIG. 7.

45 [0203] In order to confirm operation of the path search function by each of the LSRs 120R-1 to 120R-5, load distribution was performed using the network construction for experiment shown in FIG. 14, and the relationship between the traffic movement amount and the traffic convergence time with regard to the links link1 and link3 (which correspond to the entrances of the plurality of transmission paths 123-1 and 123-2, respectively) connected to the source LSR 120R-1 was evaluated as the load distribution performance by the load distribution apparatus of the present invention. Results of the evaluation are illustrated in FIGS. 15 and 16.

50 [0204] In FIG. 15, variations in time of the link utilization of the links link1 and link2 when load distribution was performed in the conditions of the input traffic of 50 Mbit/sec (93 flows), the flooding interval (OSPF flooding interval) by the OSPF of 10 sec, the congestion threshold of 30 Mbit/sec and the Hash granularity of 100 % and the link utilizations were calculated after each 10 seconds are shown in graph.

[0205] Immediately after the experiment was started, the link utilization (indicated by a void circle in the graph) of the link link1 was a little lower than 50 Mbit/sec and higher than 30 Mbit/sec of the congestion discrimination threshold value, and congestion occurred. On the other hand, the link utilization (a void triangle) of the link link3 was almost 0 Mbit/sec, and it can be seen that one-sidedness of the load appeared between the two links link1 and link3.

5 **[0206]** The load distribution by the load distribution apparatus of the present invention was started at 40 seconds after the experiment was started. The link utilization of the link link1 decreased suddenly while the link utilization of the link link3 increased suddenly, and it can be seen that the load was distributed between the two links link1 and link3. Then, at 20 seconds after the load distribution was started (at 70 seconds after the experiment was started), the traffics of the two links link1 and link3 were turned and an over-controlled state was entered once. However, while oscillations 10 were repeated, the difference became very small and the difference between the traffics of the two links link1 and link3 gradually converged. The convergence time of the traffics was 120 seconds after the load distribution was started.

10 **[0207]** On the other hand, in FIG. 16, the experiment was performed in the same conditions as those of FIG. 15 except that the Hash granularity was changed to 20 %. As a result, although the convergence time of the traffics was 200 seconds after the load distribution was started and was longer than that of the result illustrated in FIG. 15, the traffics were not turned and an over-controlled state was not entered, but the difference between the traffics of the two 15 links link1 and link3 converged asymptotically. Thus, it can be seen that distribution of the load was achieved in a stable state.

15 **[0208]** In particular, if the Hash granularity is set rather high, then although the traffics are rendered unstable through 20 over-control, load distribution can be achieved in a short time. In contrast, if the Hash granularity is set rather low, then although the time required for load distribution becomes long, load distribution can be achieved in a stable state while preventing the traffics from being rendered unstable through over-control. In other words, appropriate setting of the Hash granularity in accordance with the state of the network or the performance of the hardware allows efficient load distribution.

25 (2-7) Others

30 **[0209]** It is to be noted that, while it is described in the embodiment described above that, using MPLS (Multi-Protocol Label Switching) and using a label function of the MPLS, a route different from a routing protocol is set to perform load distribution, the present invention can be applied even if some other technique is used wherein, using an ATM (Asynchronous Transfer Mode), an FR (Frame Relay) or the like and using a protocol unique to the layer 2, a route different from the layer 3 (IP) is set to perform load distribution, or using source routing in the IP, a route different from a path of a routing protocol is set to perform load distribution.

35 **[0210]** It is also to be noted that, while a router is used as an example of the communication apparatus in the embodiment described above, the present invention is not only applied to a router but is widely applicable to other kinds of communication apparatus such as a gateway.

40 **[0211]** Further, the present invention is not limited to the embodiment described hereinabove and can be carried out in various forms without departing from the spirit of the present invention.

(2-8) Addition

45 **[0212]** As described in detail above, according to the present invention, each communication apparatus has a traffic characteristic observation function and a traffic characteristic value notification function, and a communication apparatus (source communication apparatus) which serves as a start point of load distribution can dynamically perform bandwidth management of a plurality of paths to another communication apparatus (destination communication apparatus) which serves as an end point. As a result, load distribution in accordance with free bandwidths can be performed. In particular, since a plurality of routes (transmission paths) can be set from a communication apparatus (source communication apparatus) which serves as a start point to another communication apparatus (destination communication apparatus) which serves as an end point from among communication apparatus in an IP network and distribution of load can be performed among the set routes, traffic engineering can be achieved within the network such as the Internet 50 irrespective of the network topology and without depending upon the type of the transmission paths.

55 **[0213]** The traffic characteristic collection section may smooth information regarding the collected traffic characteristic values, and this can reduce an influence of a sudden traffic variation.

60 **[0214]** Further, the traffic characteristic collection section may determine utilizations of the transmission paths connected to the communication apparatus based on the collected traffic characteristic values. This allows the utilizations of the transmission paths to be grasped particularly. In this instance, where a smoothing unit is constructed so as to smooth the utilizations determined by a utilization discriminating unit, the utilizations of the transmission paths can be grasped accurately while eliminating an influence of the traffic variation. Also where the traffic characteristic collection section has determines utilizations of the transmission paths connected to the communication apparatus based on the

traffic characteristic values smoothed by the smoothing unit, a similar effect can be achieved.

[0215] Further, if the traffic characteristic collection section determines a maximum value of the collected traffic characteristic values as a representative value of the traffic characteristic values, then transmission path selection control can be performed with a degree of certainty.

5 [0216] Further, the traffic characteristic notification section may notify (provide) information regarding the traffic characteristics making use of a packet provided by a routing protocol, and according to this construction, an existing protocol can be utilized, and consequently, it is not necessary to newly prepare a routing protocol for exclusive use, which contributes to reduction of the cost and so forth.

10 [0217] Furthermore, the traffic characteristic notification section may notify information regarding the traffic characteristics after each fixed period, and according to this construction, arithmetic load applied to an arithmetic unit such as a CPU which executes arithmetic operation and so forth can be reduced.

15 [0218] Further, the traffic characteristic notification section may notify information regarding the traffic characteristic with a message of the RSVP expanded, and according to this construct, an existing protocol can be utilized. Consequently, it is not necessary to newly prepare a routing protocol for exclusive use, which contributes to reduction of the cost and so forth.

[0219] Furthermore, the traffic characteristic collection section may collect a unit of a flux of flows for collecting traffic characteristics in a unit of a label which is used by a communication apparatus of the label switch type, and according to this construction, traffic characteristics can be collected finely, which contributes to augmentation of the collection accuracy.

20 [0220] Further, the traffic characteristic collection section provided in the source communication apparatus may determine a utilization of each of the transmission paths based on an average utilization of the transmission paths, and this can contribute to augmentation of the efficiency in use of the network.

25 [0221] Furthermore, the traffic characteristic collection section provided in the source communication apparatus may determine a utilization of each of the transmission paths based on a maximum utilization of the transmission paths, and according to this construction, transmission path selection control can be performed with certainty.

[0222] Further, the traffic characteristic collection section provided in the source communication apparatus may determine after each fixed period whether or not each of the transmission paths is congested, and according to this construction, arithmetic load applied to an arithmetic unit such as a CPU which executes arithmetic operation and so forth can be reduced.

30 [0223] Furthermore, the load calculation section provided in the source communication apparatus may calculate effective load taking a number of abandoned packets generated in the intermediate communication apparatus into consideration, and according to this construction, even if a load is not directly measured, an effective load can be estimated by a simple method. In this instance, an upper limit to the effective load can be set, and according to this construction, load estimation of an excessive amount can be avoided, and as a result, appropriate transmission path selection control can be performed.

35 [0224] Further, the decision section provided in the source communication apparatus may calculate an effective bandwidth from the effective load obtained by an effective load calculation unit and determining, regarding all of the transmission paths as one virtual pipe, a utilization of the virtual pipe, and according to this construction, it can be prevented to unnecessarily add a new transmission path and this can contribute also to effective utilization of the network.

40 [0225] Furthermore, the transmission path controlling apparatus may be constructed so as to determine whether a transmission path is to be added or deleted based on the information regarding the pipe utilization obtained by the calculation unit. This allows addition and deletion of a transmission path appropriately with the loads to the transmission paths at present taken into consideration, and distribution of the load can be achieved efficiently. In this instance, an transmission path addition-deletion discrimination unit may be constructed so as to perform discrimination of whether or not a transmission path should be added or deleted based on at least one of the pipe utilization, rate of change and integrated amount obtained by the calculation means, an appropriate path can be selected in accordance with the situation of the network or the performance of the hardware.

45 [0226] Further, the decision section can be constructed so as to compare the pipe utilization, rate of change and integrated value obtained by the calculation unit with respective reference values and perform determination of whether or not a transmission path should be added or deleted in response to a result of the comparison, and according to this construction, up-to-date optimum path selection can be performed.

50 [0227] Furthermore, the calculation unit may be constructed so as to calculate a difference value between pipe utilizations and/or a differential value of the pipe utilization as the rate of change of the pipe utilization or may be constructed so as to determine a number of times by which the pipe utilization exceeds a reference value or an integral value of the pipe utilization as the integrated amount of the pipe utilization. This allows the calculation amount to be changed in accordance with the situation of the network or the performance of the hardware and allows efficient path selection.

- [0228] Further, the decision section may select a candidate of a transmission path as an object of deletion based on the load information determined by the load calculation section, estimating the load to the other transmission paths when the transmission path of the selected candidate of deletion selected is deleted, and comparing the estimated load to the other transmission paths with a predetermined reference value and determining whether or not the transmission path of the candidate of deletion should be deleted in response to a result of the comparison. This allows appropriate selection of a transmission path to be deleted taking an influence when any transmission path is deleted into consideration. In this instance, where the information regarding the pipe utilization obtained by the calculation unit is used as a trigger for the operation, a transmission path can be deleted efficiently in response to the situation of the entire network.
- [0229] Further, the load equalization section provided in the source communication apparatus may have a move effective bandwidth calculation unit for determining an effective bandwidth to be moved with each of the transmission paths from an average effective bandwidth calculated for all of the transmission paths based on an effective bandwidth obtained from effective load obtained taking a number of abandoned packets generated in the intermediate communication apparatus into consideration, and according to this construction, equalization of load with bandwidths of transmission paths taken into consideration can be performed.
- [0230] Furthermore, the load equalization section provided in the source communication apparatus may distribute packets flowing into the source communication apparatus to the plurality of transmission paths in proportion to the effective bandwidths to be moved calculated by the move effective band calculation unit, and according to this construction, precise equalization of load can be performed.
- [0231] Further, the load equalization section provided in the source communication apparatus may be constructed so as to perform a calculation using a Hash function based on an address and distribute traffic flows at random based on a result of the calculation, and according to this construction, traffic flows can be distributed precisely by simple calculation.
- [0232] In this instance, the load equalization section may be constructed so as to perform the calculation using a CRC (cyclic redundancy check) as the Hash function or perform calculation to produce a random integer value by replacement of bit positions of an input value as the Hash function or the load equalization section may be constructed so as to additionally use a control value other than the address included in each packet as an input value for the Hash function. This allows efficient and effective distribution of traffic flows to be performed.

30

Claims

1. A transmission path controlling apparatus for use with an internet protocol network which includes a source communication apparatus, a destination communication apparatus, a plurality of transmission paths capable of being set between said source communication apparatus and said destination communication apparatus, and at least one intermediate communication apparatus each interposed in any one of said transmission paths, said transmission path controlling apparatus being provided in one of said communication apparatus, comprising:
 - a traffic characteristic collection section for collecting traffic characteristics of those of said transmission paths connected to the one communication apparatus or the other communication apparatus,
 - a traffic characteristic notification section for notifying the other communication apparatus of the traffic characteristics collected by said traffic characteristic collection section,
 - a load calculation section for calculating load based on the traffic characteristics collected by said traffic characteristic collection section,
 - a decision section for deciding whether or not a transmission path should be added or deleted based on the load information determined by said load calculation section, and
 - a load equalization section for equalizing the load determined by said load calculation section among said plurality of transmission paths.
2. A transmission path controlling apparatus for use with an internet protocol network which includes a source communication apparatus, a destination communication apparatus, a plurality of transmission paths capable of being set between said source communication apparatus and said destination communication apparatus, and at least one intermediate communication apparatus each interposed in any one of said transmission paths, said transmission path controlling apparatus being provided in said source communication apparatus, comprising:
 - a traffic characteristic collection section for collecting traffic characteristics of those of said transmission paths connected to said source communication apparatus or the other communication apparatus,
 - a load calculation section for calculating load based on the traffic characteristics collected by said traffic char-

- acteristic collection section,
a decision section for deciding addition or deletion of a transmission path based on the load information determined by said load calculation section, and
a load equalization section for equalizing the load determined by said load calculation section among said plurality of transmission paths.
- 5 3. A transmission path controlling apparatus for use with an internet protocol network which includes a source communication apparatus, a destination communication apparatus, a plurality of transmission paths capable of being set between said source communication apparatus and said destination communication apparatus, and at least one intermediate communication apparatus each interposed in any one of said transmission paths, said transmission path controlling apparatus being provided in any one of the communication apparatus other than said source communication apparatus, comprising:
- 10 a traffic characteristic collection section for collecting traffic characteristic of the transmission paths connected to the communication apparatus, and
15 a traffic characteristic notification section for notifying said source communication apparatus of the traffic characteristic collected by said traffic characteristic collection section.
- 20 4. The transmission path controlling apparatus of any one of claims 1 to 3, further comprising:
 in said traffic characteristic collection section, means for smoothing the collected traffic characteristic values.
- 25 5. The transmission path controlling apparatus of claim 4, further comprising:
 in said traffic characteristic collection section, means for calculating utilizations of the transmission paths connected to the communication apparatus based on the collected traffic characteristic values.
- 30 6. The transmission path controlling apparatus of claim 5, wherein said smoothing means of said traffic characteristic collection section is constructed so as to smooth the utilizations calculated by said utilization calculating means.
- 35 7. The transmission path controlling apparatus of claim 4, wherein said smoothing means of said traffic characteristic collection section is constructed so as to smooth the collected traffic characteristic values, and said traffic characteristic collection section has means for calculating utilizations of the transmission paths connected to the communication apparatus based on the traffic characteristic values smoothed by said smoothing means.
- 40 8. The transmission path controlling apparatus of any one of claims 1 to 3, further comprising:
 in said traffic characteristic collection section, means for determining a maximum value of the collected traffic characteristic values as a representative value of the traffic characteristic values.
- 45 9. The transmission path controlling apparatus of claim 1 or 3, further comprising:
 in said traffic characteristic notification section, means for notifying of information regarding the traffic characteristics making use of a packet provided by a routing protocol.
- 50 10. The transmission path controlling apparatus of any one of claims 1, 3 and 9, further comprising:
 in said traffic characteristic notification section, means for notifying of information regarding the traffic characteristics at periodical intervals.
- 55 11. The transmission path controlling apparatus of claim 1 or 3, further comprising:
 in said traffic characteristic notification section, means for notifying of information regarding the traffic characteristic with a message of the RSVP expanded.
- 60 12. The transmission path controlling apparatus of any one of claims 1 to 3, further comprising:
 in said traffic characteristic collection section, means for collecting a unit of a flux of flows for collecting traffic characteristics in a unit of a label which is used by a communication apparatus of the label switch type.
- 65 13. The transmission path controlling apparatus of claim 1 or 2, further comprising:
 in said traffic characteristic collection section provided in said source communication apparatus, means for calculating a utilization of each of the transmission paths connected to said source communication apparatus based on an average utilization of the transmission paths connected to said intermediate communication apparatus collected from said intermediate communication apparatus.

14. The transmission path controlling apparatus of claim 13, further comprising:
in said traffic characteristic collection section provided in said source communication apparatus, means for calculating a utilization of each of the transmission paths connected to said source communication apparatus based on a maximum utilization of the transmission paths connected to said intermediate communication apparatus collected from said intermediate communication apparatus.
- 5
15. The transmission path controlling apparatus of claim 1 or 2, further comprising:
in said traffic characteristic collection section provided in said source communication apparatus, means for discriminating at periodical intervals whether or not each of said transmission paths is congested.
- 10
16. The transmission path controlling apparatus of claim 1 or 2, further comprising:
in said load calculation section provided in said source communication apparatus, means for calculating effective load taking a number of abandoned packets generated in said intermediate communication apparatus into consideration.
- 15
17. The transmission path controlling apparatus of claim 16, wherein said effective load calculation means is constructed so as to set an upper limit to the effective load when the effective load is to be calculated.
- 20
18. The transmission path controlling apparatus of claim 1 or 2, further comprising:
in said load calculation section provided in said source communication apparatus, means for calculating effective load taking a number of abandoned packages generated in said intermediate communication apparatus into consideration, and
said decision section provided in said source communication apparatus has means for calculating an effective bandwidth from the effective load obtained by said effective load calculation means and for determining, regarding all of said transmission paths as one virtual pipe, a utilization of the virtual pipe.
- 25
19. The transmission path controlling apparatus of claim 1 or 2, further comprising:
in said load calculation section provided in said source communication apparatus, means for calculating effective load taking a number of abandoned packages generated in said intermediate communication apparatus into consideration, and
said decision section provided in said source communication apparatus having
30
35
means for calculating an effective bandwidth from the effective load calculated by said effective load calculation means and for determining, regarding all of said transmission paths as one virtual pipe, a utilization of the virtual pipe, and
means for comparing the pipe utilization obtained by said pipe utilization calculation means with a set reference value and for deciding whether a transmission path is to be added or deleted in accordance with a result of the comparison.
- 40
20. The transmission path controlling apparatus of claim 19, wherein
said calculation means of said decision section is constructed so as to calculate at least one of the pipe utilization, a rate of change of the pipe utilization per time and an integrated amount of the pipe utilization in a fixed time, and
45
said transmission path addition-deletion decision means is constructed so as to perform decision of whether or not a transmission path should be added or deleted based on at least one of the pipe utilization, rate of change and integrated amount obtained by said calculation means.
- 50
21. The transmission path controlling apparatus of claim 20, wherein said decision section provided in said source communication apparatus is constructed so as to perform comparison of at least one of a set of the pipe utilization obtained by said calculation means and a first reference value, another set of the rate of change obtained by said calculation means and a second reference value and a further set of the integrated value obtained by said calculation means and a third reference value, and perform decision of whether or not a transmission path should be added or deleted in response to a result of the comparison.
- 55
22. The transmission path controlling apparatus of claim 21, wherein said decision section provided in said source communication apparatus is constructed so as to add a transmission path when the pipe utilization obtained by

EP 1 130 849 A2

said pipe utilization calculation means is higher than a reference value for addition of a path.

23. The transmission path controlling apparatus of claim 21, wherein said decision section provided in said source communication apparatus is constructed so as to delete a transmission path having a low load when the pipe utilization obtained by said pipe utilization calculation means is lower than a reference value for deletion of a path.

5 24. The transmission path controlling apparatus of any one of claims 19 to 23, wherein said calculation means is constructed so as to calculate at least one of a difference value between two pipe utilizations obtained after a very small interval and a differential value of the pipe utilization at a certain point of time as the rate of change of the pipe utilization per time.

10 25. The transmission path controlling apparatus of any one of claims 19 to 23, wherein said calculation means is constructed so as to calculate at least one of a number of times by which the pipe utilization exceeds a certain reference value within a fixed period of time and an integral value of the pipe utilization within the fixed time as the integrated amount of the pipe utilization within the fixed period of time.

15 26. The transmission path controlling apparatus of claim 1 or 2, wherein

20 said decision section provided in said source communication apparatus includes
means for selecting a candidate of a transmission path as an object of deletion based on the load information determined by said load calculation section,
means for estimating the load to the other transmission paths when the transmission path of the candidate of deletion selected by said deletion candidate selection means is deleted, and
25 means for comparing the load to the other transmission paths estimated by said load estimation means with a predetermined reference value and determining whether or not the transmission path of the candidate of deletion should be deleted in response to a result of the comparison.

27. The transmission path controlling apparatus of claim 26, further comprising:

30 in said load calculation section provided in said source communication apparatus, means for calculating an effective load taking a number of abandoned packets generated in said intermediate communication apparatus into consideration, and
said decision section provided in said source communication apparatus further includes
35 calculation means for calculating an effective bandwidth from the effective load obtained by said effective load calculation means and determining, regarding all of said transmission paths as one virtual pipe, information regarding a utilization of the virtual pipe, and
trigger means for triggering said selection means, estimation means and determination means to operate based on the information regarding the pipe utilization obtained by said calculation means.

40 28. The transmission path controlling apparatus of claim 1 or 2, further comprising:

45 in said load equalization section provided in said source communication apparatus, move effective bandwidth calculation means for determining an effective bandwidth to be moved with each of said transmission paths from an average effective bandwidth calculated for all of said transmission paths based on an effective bandwidth obtained from effective load obtained taking a number of abandoned packets generated in said intermediate communication apparatus into consideration.

29. The transmission path controlling apparatus of claim 28, further comprising:

50 in said load equalization section provided in said source communication apparatus, means for distributing packets flowing into said source communication apparatus to said plurality of transmission paths in proportion to the effective bandwidths to be moved calculated by said move effective band calculation means.

30. The transmission path controlling apparatus of claim 29, wherein said load equalization section provided in said source communication apparatus is constructed so as to perform a calculation using a Hash function based on an address and distribute traffic flows at random based on a result of the calculation.

55 31. The transmission path controlling apparatus of claim 30, wherein said load equalization section provided in said source communication apparatus performs the calculation using a CRC (cyclic redundancy check) as the Hash function.

32. The transmission path controlling apparatus of claim 30, wherein said load equalization section provided in said source communication apparatus performs at least one of replacement of positions and logical operation for bit values of an input value as the calculation by the Hash function to produce a random integer value.
- 5 33. The transmission path controlling apparatus of claim 30, wherein said load equalization section provided in said source communication apparatus additionally uses, as an input value for the Hash function, a control value other than the address included in each packet.
- 10 34. A transmission path controlling method for an internet protocol network which includes a source communication apparatus, a destination communication apparatus, a plurality of transmission paths capable of being set between said source communication apparatus and said destination communication apparatus, and at least one intermediate communication apparatus each interposed in any one of said transmission paths, the method comprising:
- 15 (a) collecting traffic characteristics of those of said transmission paths connected to other communication apparatus, and
 (b) notifying said source communication apparatus of the traffic characteristics collected by step (a),
 steps (a) and (b) being performed by a communication apparatus other than said source communication apparatus of said internet protocol network, and
20 (c) collecting traffic characteristics of those of said transmission paths connected to said source communication apparatus,
 (d) calculating load based on one or both of the traffic characteristics collected by the traffic characteristic collection step and the traffic characteristics obtained from the other communication apparatus,
 (e) deciding whether or not a transmission path should be added or deleted based on the load information determined by the load calculation step, and
25 (f) equalizing the load determined by the load calculation step among said plurality of transmission paths,
 steps (d), (e), and (f) being performed by said source communication apparatus.
- 30 35. A medium on which a transmission path controlling program for use with an internet protocol network which includes a source communication apparatus, a destination communication apparatus, a plurality of transmission paths capable of being set between said source communication apparatus and said destination communication apparatus, and at least one intermediate communication apparatus each interposed in any one of said transmission paths is recorded, said transmission path controlling program being used by one of said communication apparatus, said transmission path controlling program causing a computer to function as
- 35 traffic characteristic collection means for collecting traffic characteristics of those of said transmission paths connected to the communication apparatus or the other communication apparatus,
 traffic characteristic notification means for notifying the other communication apparatus of the traffic characteristics collected by said traffic characteristic collection means,
 load calculation means for calculating load based on the traffic characteristics collected by said traffic characteristic collection means,
40 decision means for deciding whether or not a transmission path should be added or deleted based on the load information determined by said load calculation means, and
 load equalization means for equalizing the load determined by said load calculation means among said plurality of transmission paths.
- 45 36. A network load balancing apparatus which balances the traffic load among transmission paths set between an ingress node and an egress node in a network having a plurality of network nodes, said network having a plurality of transmission paths capable of being set between said ingress node and said egress node, comprising:
- 50 a traffic characteristic collection section for determining traffic characteristics of said transmission paths set between said ingress node and said egress node,
 a traffic characteristic notification section for notifying network nodes of results of said traffic characteristic collection section,
 a load calculation section for calculating traffic load based on the traffic characteristics collected by said traffic characteristic collection section, and
55 a means for causing transmission of a packet on a transmission path set between said ingress node and said egress node depending upon said traffic load calculation.

EP 1 130 849 A2

37. The network load balancing apparatus of claim 36, further comprising:
a decision section for deciding whether or not a transmission path should be added or deleted between said ingress node and said egress node based on the traffic load information determined by said load calculation section.
- 5 38. The network load balancing apparatus of claim 36, further comprising:
a load equalization section for equalizing the load determined by said load calculation section among said set transmission paths.
- 10 39. The network load balancing apparatus of claim 36, further comprising:
means in said traffic characteristic notification section for providing said results of the information regarding the traffic characteristics periodically and by making use of a packet provided by a routing protocol.
- 15 40. The network load balancing apparatus of claim 36, wherein
said traffic characteristic collection section calculates occupancy of the set transmission paths based on the collected traffic characteristic values.
- 20 41. The network load balancing apparatus of claim 40, wherein
said traffic characteristic collection section smooths the calculated occupancy of the set transmission paths.
- 25 42. The network load balancing apparatus of claim 40, wherein
said traffic characteristic collection section smooths the collected traffic characteristic values, and said traffic characteristic collection section calculates occupancy of the set transmission paths based on the smoothed collected traffic characteristic values.
- 30 43. The network load balancing apparatus of claim 36, further comprising:
means in said traffic characteristic notification section for collecting a unit of flux of flows for thereby collecting traffic characteristics in a unit of a label utilized by a router of a label switch type.
- 35 44. A network load balancing apparatus which balances the traffic load among transmission links which link node devices in a network, and where at least one transmission link defines a transmission path between an ingress node and an egress node, and said network having a plurality of transmission paths capable of being set between said ingress node and said egress node, comprising:
a traffic characteristic collection section for collecting traffic characteristics of said transmission links being portions of transmission paths set between said ingress node and said egress node,
a traffic characteristic notification section for providing results of said traffic characteristic collection section,
a load calculation section for calculating traffic load based on the traffic characteristics collected by said traffic characteristic collection section, and
a means for causing transmission of a packet on a transmission path set between said ingress node and said egress node depending upon said traffic load calculation.
- 40 45. The network load balancing apparatus of claim 44, further comprising:
a decision section for deciding whether or not a transmission path should be added or deleted between said ingress node and said egress node based on the load information determined by said load calculation section.
- 45 46. The network load balancing apparatus of claim 44, further comprising:
a load equalization section for equalizing the load determined by said load calculation section among said set transmission paths.
- 50 47. The network load balancing apparatus of claim 44, further comprising:
a means provided in said traffic characteristic collection section for calculating an occupancy of each transmission path among said set transmission paths based on an average occupancy of transmission links comprising said each transmission path.
- 55 48. The network load balancing apparatus of claim 44, further comprising:
a means provided in said traffic characteristic collection section for calculating an occupancy of each transmission path among said set transmission paths based on a maximum occupancy of transmission links comprising said each transmission path.

EP 1 130 849 A2

49. A transmission path controlling apparatus for use with an internet protocol network which includes a source router, a destination router, a plurality of transmission paths capable of being set between said source router and said destination router, said internet protocol network having a plurality of network devices including said source and destination routers, said transmission path controlling apparatus comprising:

- 5 a traffic characteristic collection section for collecting traffic characteristics of those of said transmission paths set between said source router and said destination router,
10 a traffic characteristic notification section for notifying network devices of results from said traffic characteristic collection section,
15 a load calculation section for calculating load based on the traffic characteristics collected by said traffic characteristic collection section, and
20 a means for causing transmission of a packet on a transmission path set between said source router and said destination router depending upon said traffic load calculation.

25 50. The transmission path controlling apparatus of claim 49, further comprising:
 a decision section for deciding whether or not a transmission path should be added or deleted based on the load determined by said load calculation section.

20 51. The transmission path controlling apparatus of claim 49, further comprising:
 a load equalization section for equalizing the load determined by said load calculation section among said set transmission paths.

25 52. The transmission path controlling apparatus of claim 49, further comprising:
 means for smoothing the collected traffic characteristics.

30 53. The transmission path controlling apparatus of claim 52, wherein said traffic characteristic collection section calculates the occupancy of those of said transmission paths set between said source router and said destination router based on the smoothed traffic characteristics.

35 54. The transmission path controlling apparatus of claim 49, further comprising:
 means for determining a maximum value of the collected traffic characteristics as a representative value of the traffic characteristic values.

40 55. The transmission path controlling apparatus of claim 49, further comprising:
 means for periodically notifying information regarding the traffic characteristics making use of a packet provided by a routing protocol.

45 56. The transmission path controlling apparatus of claim 49, further comprising:
 means for notifying information regarding the traffic characteristic with a message of RSVP expanded protocol.

50 57. The transmission path controlling apparatus of claim 49, further comprising:
 means for collecting a unit of a flux of flows thereby collecting traffic characteristics in a unit of a label which is used by a router of a label switch type.

55 58. The transmission path controlling apparatus of claim 49, further comprising:
 means for calculating an occupancy of each transmission path connected to said source router based on an average occupancy of the transmission paths connected to an intermediate router collected from said intermediate router, wherein said intermediate router may be interposed in any one of said transmission paths.

55 59. The transmission path controlling apparatus of claim 49, further comprising:
 means for calculating an occupancy of each of the transmission paths connected to said source router based on a maximum occupancy of the transmission paths connected to an intermediate router collected from said intermediate router, wherein said intermediate router may be interposed in any one of said transmission paths.

55 60. The transmission path controlling apparatus of claim 49, further comprising:
 means for determining periodically whether or not each of said transmission paths is congested.

EP 1 130 849 A2

61. The transmission path controlling apparatus of claim 49, further comprising:
means provided in said load calculation section for calculating effective load taking a number of abandoned packets generated in an intermediate router into consideration, wherein said intermediate router may be interposed in any one of said transmission paths.
- 5 62. The transmission path controlling apparatus of claim 61, wherein
 said means is constructed so as to set an upper limit to the effective load when the effective load is to be calculated.
- 10 63. The transmission path controlling apparatus of claim 50, further comprising:
 means provided in said load calculation section for calculating effective load taking a number of abandoned packages generated in an intermediate router into consideration, and
15 means provided in said decision section for calculating an effective bandwidth from the effective load obtained by said effective load calculation means and determining, regarding all of said transmission paths as one virtual pipe, an occupancy of the virtual pipe.
- 20 64. The transmission path controlling apparatus of claim 50, further comprising:
 means for calculating effective load by considering a number of abandoned packages generated in an intermediate router, and
 said decision section having a pipe utilization calculation unit for calculating an effective bandwidth from the effective load calculated by said effective load calculation means and determining, regarding all of said transmission paths as one virtual pipe, a utilization of the virtual pipe, and
25 means for comparing the pipe utilization obtained by said pipe utilization calculation unit with a set reference value and determining whether a transmission path is to be added or deleted in accordance with a result of the comparison.
- 30 65. The transmission path controlling apparatus of claim 64, wherein said decision section is constructed so as to add a transmission path when the pipe utilization calculation obtained by said pipe utilization calculation unit is higher than a reference value for addition of a path.
- 35 66. The transmission path controlling apparatus of claim 64, wherein said decision section is constructed so as to delete a transmission path having a low load when the pipe utilization obtained by said pipe utilization calculation unit is lower than a reference value for deletion of a path.
- 40 67. The transmission path controlling apparatus of claim 51, further comprising:
 means for determining an effective bandwidth to be moved among each of said transmission paths from an average effective bandwidth calculated for all of said transmission paths based on an effective bandwidth obtained from effective load obtained taking a number of abandoned packets generated in an intermediate router into consideration.
- 45 68. The transmission path controlling apparatus of claim 67, further comprising:
 means for distributing packets flowing into said source router to said plurality of transmission paths in proportion to the effective bandwidths to be moved calculated by said move effective band calculation means.
- 50 69. The transmission path controlling apparatus of claim 68, wherein said load equalization section performs a calculation using a hash function based on an address and distributes traffic flows at random based on a result of the calculation.
- 55 70. A transmission path controlling method for a network which includes a source router, a destination router, and a plurality of transmission paths capable of being set between said source router and said destination router, comprising the steps of:
 collecting traffic characteristics of transmission paths set between said source router and said destination router,
 calculating load of transmission paths set between said source router and said destination router based results of the traffic characteristics collection step, and

transmitting a packet on a transmission path set between said ingress node and said egress node depending upon said traffic load calculation.

- 5 **71.** The transmission path controlling method of claim 70, further comprising the step of:
determining whether or not a transmission path should be added or deleted based on the load information determined by the load calculation step.
- 10 **72.** The transmission path controlling method of claim 70, further comprising the step of:
equalizing the load determined by the load calculation step among said plurality of transmission paths.
- 15 **73.** The transmission path controlling method of claim 70, further comprising the step of:
providing the results of the traffic characteristic collection step to node devices of the network.
- 20 **74.** A transmission path controlling method for an internet protocol network which includes a source router, a destination router, a plurality of transmission paths capable of being set between said source router and said destination router, and an intermediate router interposed in any one of said transmission paths, comprising the steps of:
collecting traffic characteristics of transmission paths connected to at least one of intermediate router and the destination router, and
notifying said source router of the results of the traffic characteristics collection step.
- 25 **75.** A computer-readable medium having computer executable instructions for use in an internet protocol network which includes a source router, a destination router, a plurality of transmission paths capable of being set between said source router and said destination router, said computer executable instructions being used by a router, and for performing the steps comprising:
collecting traffic characteristics of transmission paths set between said source router and said destination router, and
notifying other routers of the traffic characteristics collected.
- 30 **76.** A computer-readable medium of claim 75, having computer executable instructions for performing the further steps comprising:
calculating load based on the traffic characteristics collected, and
transmitting a packet based on the results of the calculated load.
- 35 **77.** A computer-readable medium of claim 75, having computer executable instructions for performing the further steps comprising:
determining whether or not a transmission path should be added or deleted based on the load information determined by said load calculation means, and
load equalization means for equalizing the load determined by said load calculation means among said plurality of transmission paths.
- 40 **78.** A label switch router in a Multi-Protocol Label Switch network, comprising:
a traffic characteristic collection section for collecting for each port thereof, traffic characteristic values including an output packet amount from an output port and a logical bandwidth of said output port as traffic characteristic values; and
50 a traffic characteristic notification section for notifying network nodes of results from said traffic characteristic collection section.
- 45 **79.** The label switch router of claim 78, further comprising:
means for processing the obtained traffic characteristic values by calculating a transmission line utilization value CUTY from the obtained traffic characteristic values in accordance with the following expression
CUTY = output packet amount from an output port/logical bandwidth of said output port.

80. The label switch router of claim 79, further comprising:
 means for processing the obtained traffic characteristic values by performing smoothing of the collected traffic characteristic values.

5 81. The label switch router of claim 80, wherein said smoothing processing is performed by calculating a value MUTY each time an observation value is obtained, wherein

$$\text{MUTY} = \alpha \times \text{CUTY} + (1 - \alpha) \times \text{MUTY}$$

10 where α is a smoothing coefficient.

82. The label switch router of claim 81, further comprising:
 means for determining a maximum value of collected characteristic values as a representative value of the traffic characteristic values, wherein the maximum value of the observation values is placed as a working value MUTY.

15 83. The label switch router of claim 82, wherein said traffic characteristic notification section distributes said traffic characteristic values including MUTY periodically utilizing a routing protocol.

20 84. The label switch router of claim 78, wherein said traffic characteristic values including an abandoned packet number NLOSS at said output port in a time period, said label switch router further comprising:

25 means for processing the obtained traffic characteristic values by calculating a total number TLOSS of abandoned packets in accordance with the following expression.

$$\text{TLOSS}(n) = \text{TLOSS}(n-1) + \text{NLOSS}$$

30 where TLOSS(n) is the total number of abandoned packets during time period n, TLOSS(n-1) is the total number of abandoned packets in a prior time period.

35 85. The label switch router of claim 84, wherein said traffic characteristic notification section distributes said traffic characteristic values including TLOSS(n) periodically utilizing a routing protocol.

36 86. The label switch router of claim 85, wherein said distribution utilizes a message of an RSVP expanded protocol to notify another label switch router of the information.

40 87. The label switch router of claim 78, further comprising:
 a traffic characteristic collection section for receiving traffic characteristics from another label switch router.

45 88. The label switch router of claim 81, further comprising:

a traffic characteristic collection section for receiving traffic characteristics from another label switch router,
 a load calculation section for calculating traffic load based on at least one of traffic characteristics collected
 from said traffic characteristic collection section and received traffic characteristics from another label switch
 router, and
 a packet transfer section for transferring a packet based on the results of said load calculation section.

50 89. The label switch router of claim 88, wherein
 said load calculation section further calculates a congestion state of individual paths from said label switch router to another label switch router where a plurality of paths are set between said label switch router to said another label switch router, and each path may extend through a plurality of repeating label switch routers and said path is composed of links between said repeating label switch routers, said links connecting said repeating label switch routers to form said path, the congestion state for each path is determined, from an occupancy of each path, collected from each of said repeating label switch routers of each path, based on an average occupancy of each link connected to each repeating label switch router.

EP 1 130 849 A2

90. The label switch router of claim 89, wherein
the occupancy of a link j of a path i are averaged and may be represented in such an expression as

5 $\rho_{\text{path } i} = \text{Average}(\text{MUTY}(\text{link } j, \text{path } i)).$

91. The label switch router of claim 90, wherein
a maximum occupancy of each link connected to each repeating label switch router is calculated where a maximum value of occupancy of all links of a path i is adopted and may be represented in the expression

10 $\rho_{\text{path } i} = \text{Max}(\text{MUTY}(\text{link } j, \text{path } i)).$

92. The label switch router of claim 84, further comprising:

15 a traffic characteristic collection section for receiving traffic characteristics from another label switch router,
a load calculation section for calculating traffic load based on at least one of traffic characteristics collected
from said traffic characteristic collection section and received traffic characteristics from another label switch
router, and
20 a packet transfer section for transferring a packet based on the results of said load calculation section.

93. The label switch router of claim 92, wherein

25 a plurality of paths are set between said label switch router to another label switch router, and each route
extends through a plurality of repeating label switch routers and is composed of links between said repeating label
switch routers, the congestion state for each route is determined from an occupancy of each path, collected from
each of said repeating label switch routers of each path, based on an average occupancy of each link connected
to each repeating label switch router.

94. The label switch router of claim 93, wherein

30 a sum total TLOSS path i of abandoned packets of said path i is determined by the expression

$$\text{TLOSS path } i = \sum \text{TLOSS link } j$$

35 where i is the number of a label switch path and j is a link number.

95. The label switch router of claim 94, wherein

calculation of an effective load $\rho_{\text{effective}}$ of path i is calculated by the expression

40 $\rho_{\text{effective path } i} = \rho_{\text{path } i} \times f(\text{TLOSS path } i)$

where $\rho_{\text{path } i}$ is the load of path i and $f(\text{TLOSS path } i)$ is a function regarding TLOSS path i.

- 45 96. The label switch router of claim 95, wherein

if packet loss occurs, then the load $\rho_{\text{path } i}$ is corrected with a predetermined function $f(\text{TLOSS path } i)$ according to the expression

50 $\rho_{\text{effective path } i} = \text{Min}(\rho_{\text{effective path } i}, \rho_{\text{ceiling}})$

where $f(\text{TLOSS path } i)$ is a function regarding TLOSS path i, and ρ_{ceiling} is an upper limit load value to the effective load $\rho_{\text{effective path } i}$.

- 55 97. The label switch router of claim 94, wherein each route extends through a plurality of repeating label switch routers, load based on traffic characteristics may be calculated utilizing a number TLOSS path i of abandoned packets generated in each repeating label switch router,

EP 1 130 849 A2

the effective load $\rho_{\text{effective path } i}$ in accordance with

$$\text{effective load to ith path} = \text{load to ith path} \times f(\text{abandoned packet number})$$

5

where $f(\text{abandoned packet number})$ is a function wherein the abandoned packet number is a variable.

98. The label switch router of claim 94, further comprising:

10

a load equalization section for adjusting the load periodically where all paths are regarded as a single virtual pipe, and a load to said virtual pipe is determined by means for calculating effective bandwidths from effective loads obtained by said effective load calculation section and determining a utilization of the virtual pipe.

99. The label switch router of claim 98, wherein

15

said utilization of the virtual pipe is determined in accordance with the expression

20

$$\text{utilization } \rho_{\text{effective all}} \text{ when all paths are regarded as one virtual pipe} = \frac{\sum (\rho_{\text{effective path } i} \times \text{LBW path } i)}{\sum \text{LBW path } i}$$

where $\rho_{\text{effective path } i}$ is the effective load of the path i , and $\text{LBW path } i$ is the logical bandwidth of the path i .

100. The label switch router of claim 99, wherein

25

said utilization of the virtual pipe is compared with a set reference value ρ_{offer1} , and when the obtained value of the utilization is higher than ρ_{offer1} , a new transmission path is added.

101. The label switch router of claim 99, wherein

said utilization of the virtual pipe is compared with a set reference value ρ_{offer2} , and if the obtained value of the utilization is lower than ρ_{offer2} , then a path is deleted.

30

35

40

45

50

55

FIG.1(a)

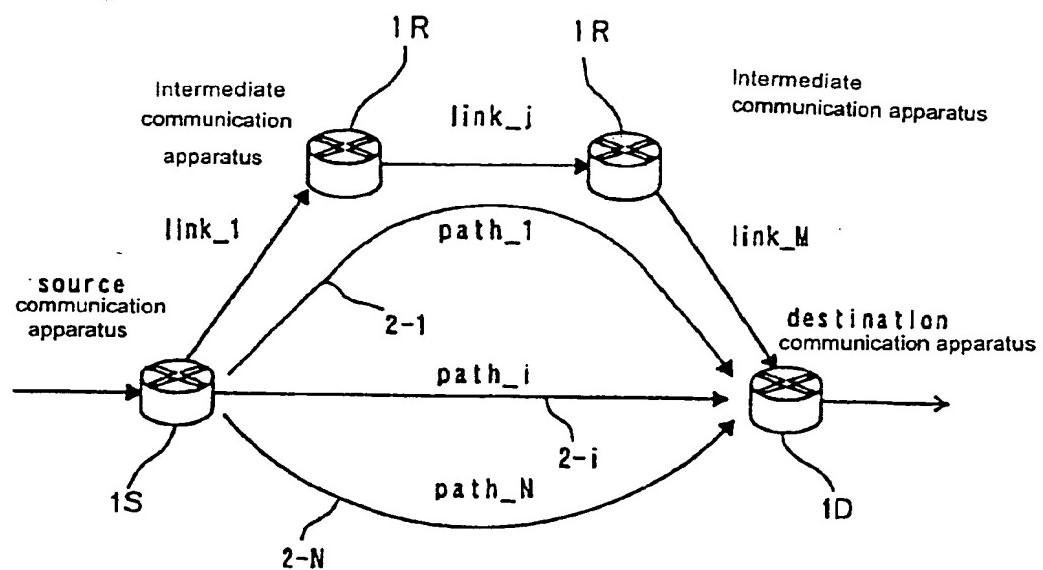


FIG.1(b)

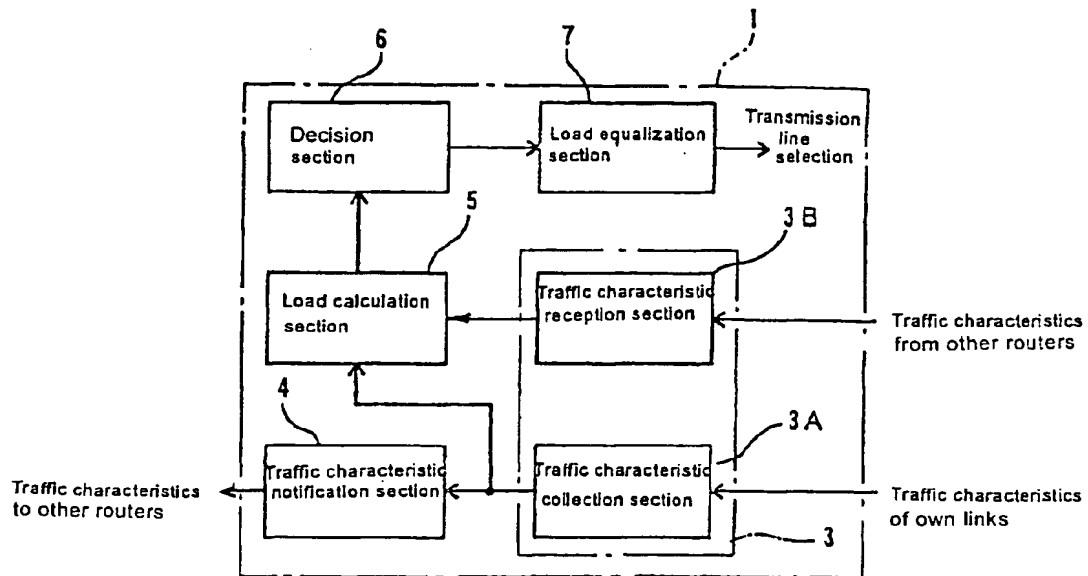


FIG.1(c)

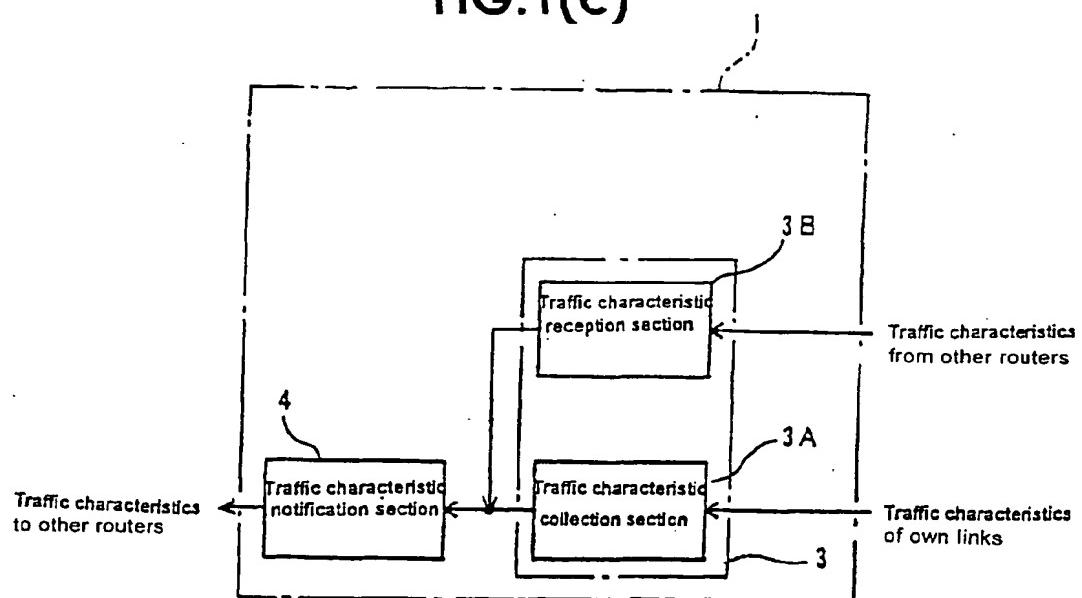


FIG.2(a)

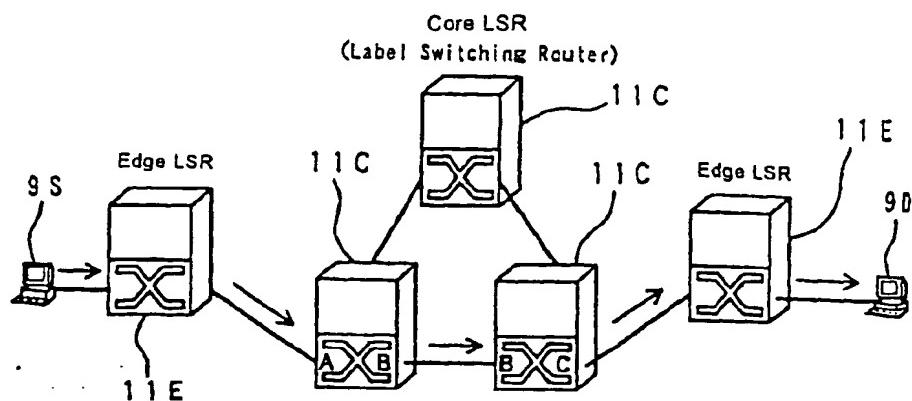


FIG.2(b)

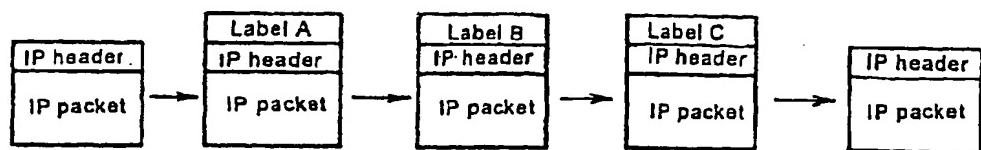


FIG.3

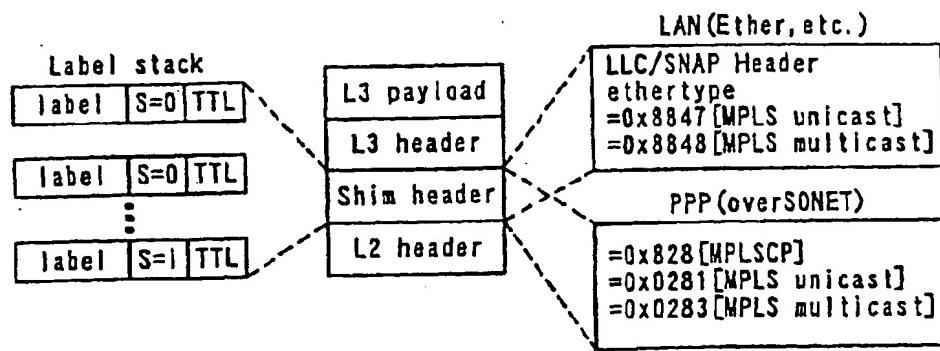


FIG.4

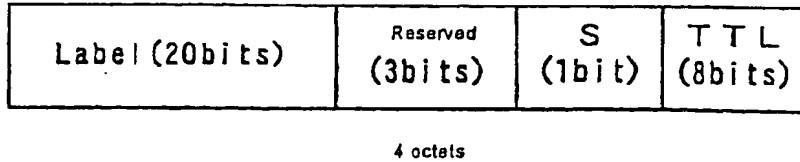


FIG.5

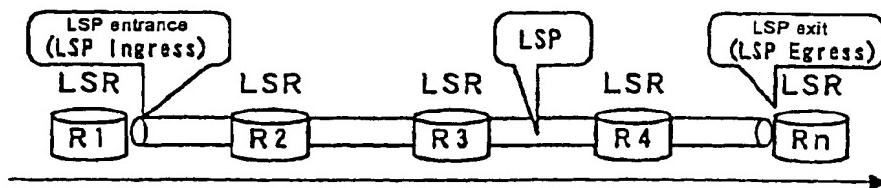


FIG.6

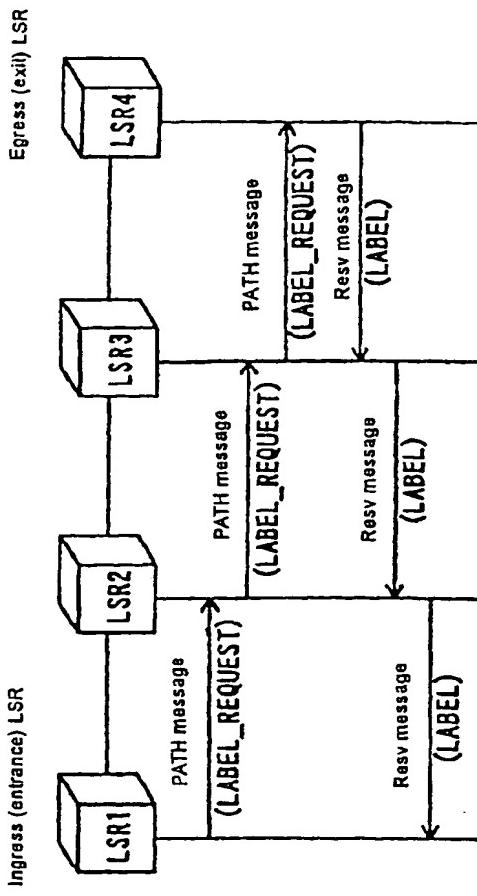


FIG.7

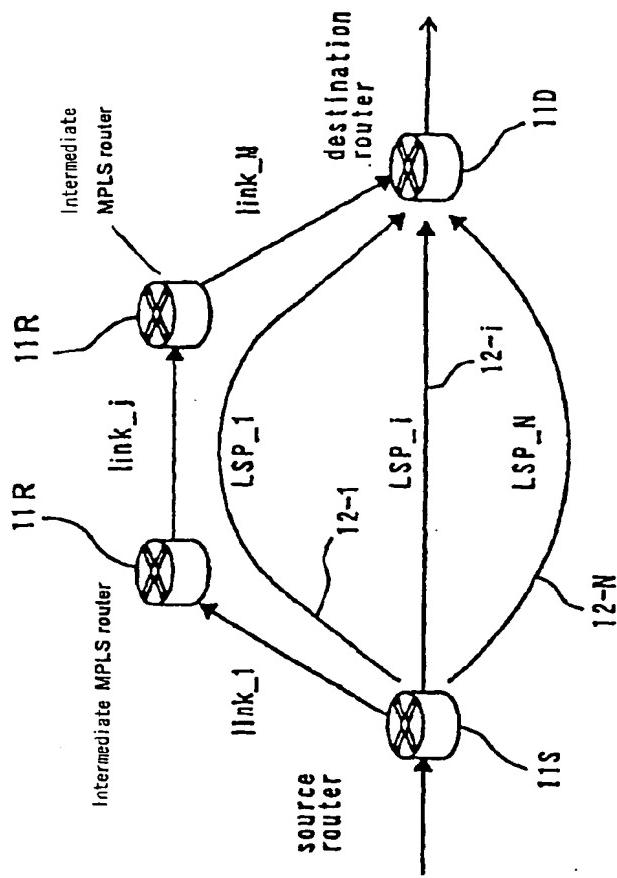


FIG.8

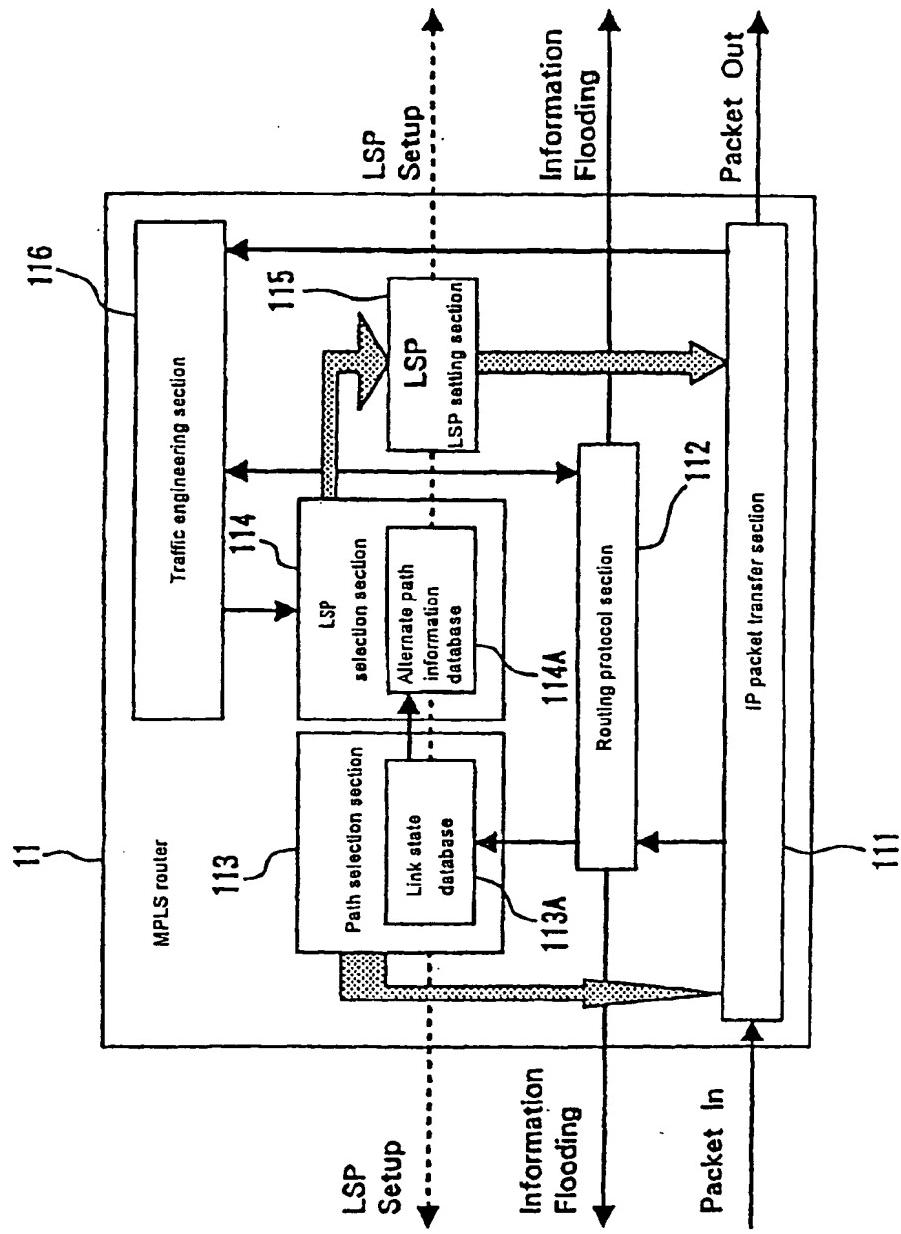


FIG. 9

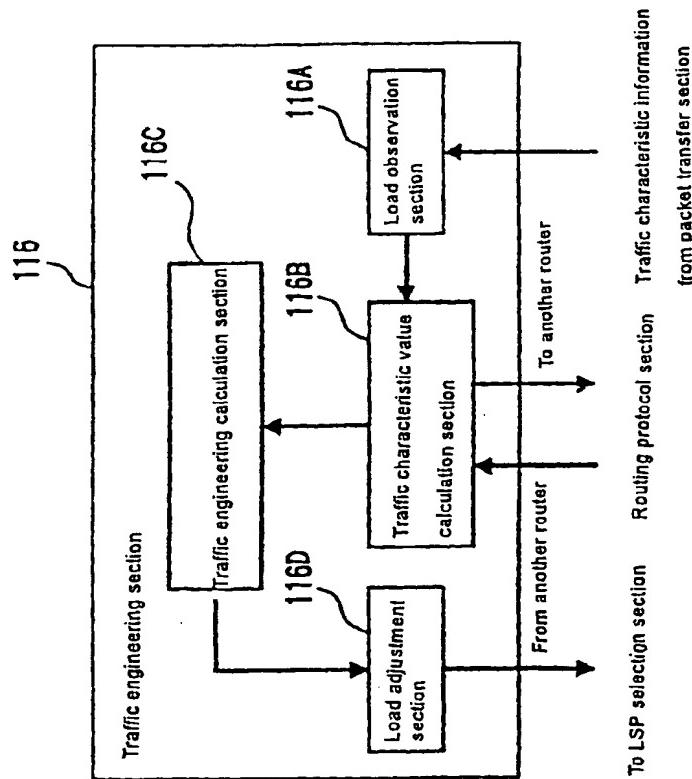


FIG.10

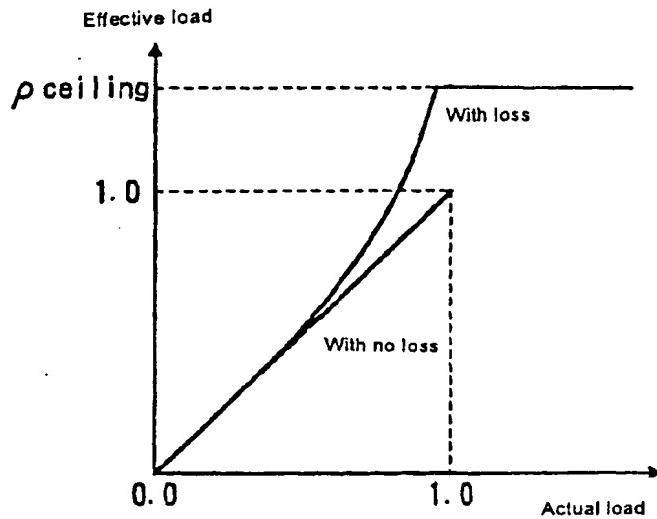


FIG.11

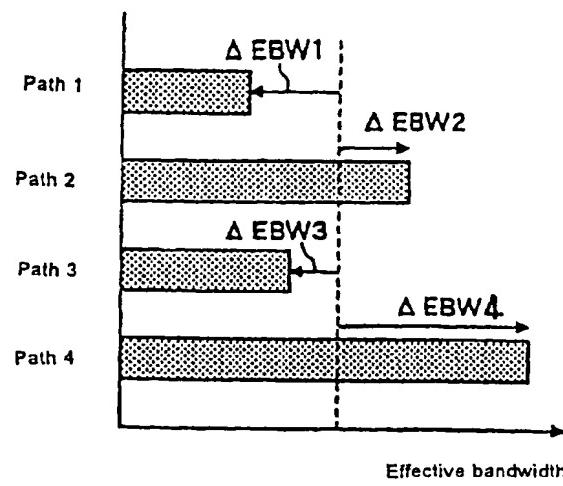


FIG.12

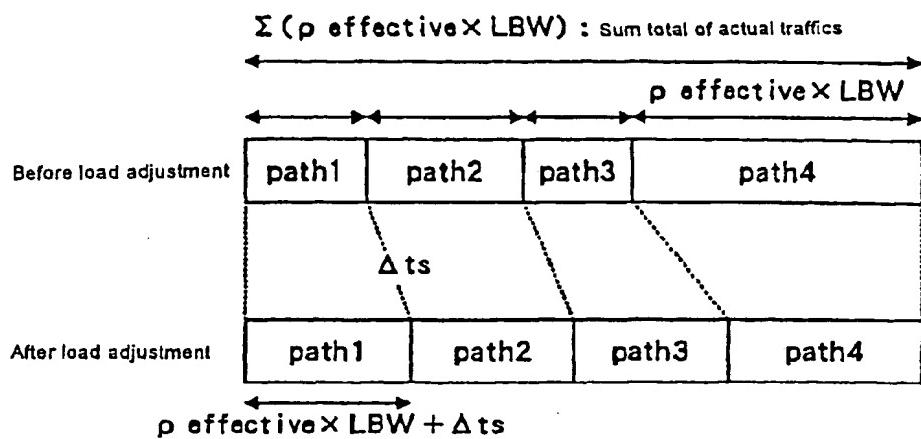


FIG.13

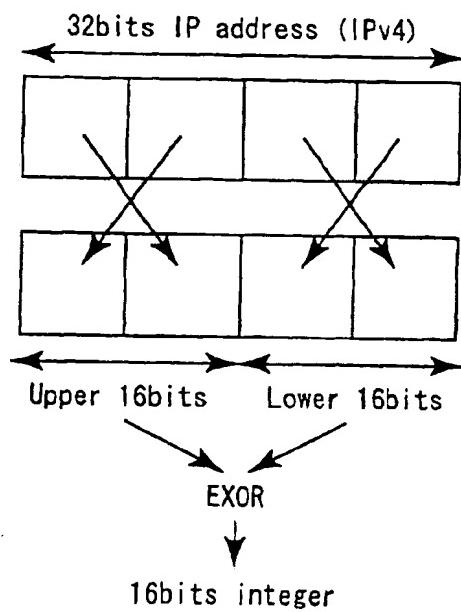


FIG.14

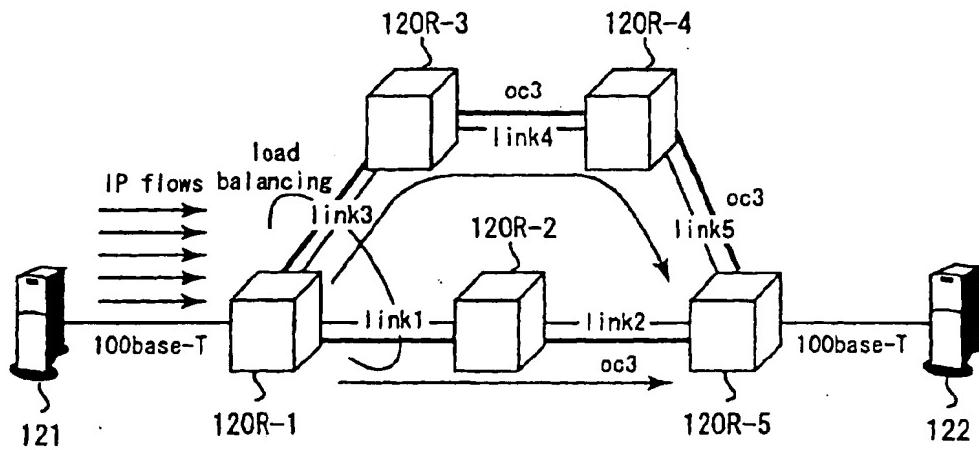


FIG.15

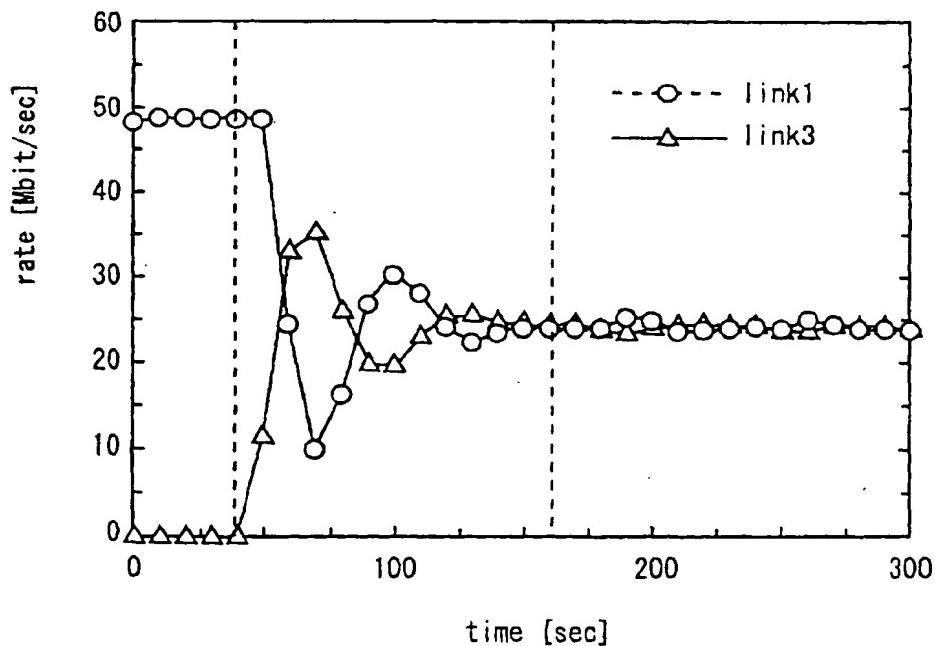


FIG.16

